SCHOOL OF EDUCATION
LIBRARY
OCT 18 1949

The Journal of Experimental Education

A periodical report of scientific investigations relating to child development, curriculum, learning, teaching, supervision, measurements, statistics, and experimental techniques.

Volume XVIII | September, 1949 | Number 1

STUDIES IN TEACHER EDUCATION

A Follow-Up of Some Earlier Wisconsin Studies of Teaching Ability: Willard Joseph Brandt... 1
A Study of Some Factors in Pupil Growth: Kenneth Peter Riesch.....31
A Study of the Distribution of Teacher Approval and Disapproval Among Sixth-Grade Pupils: Albert F. deGroat and George G. Thompson.....57

$5.00 a Year | Published Quarterly | $1.50 a Copy

Published by Dunbar Publications, Inc., 114 S. Carroll Street, Madison 3, Wisconsin.
Entered as second-class matter October 17, 1932 at the post office at Madison, Wisconsin, under the act of March 3, 1879.

EDITORIAL BOARD

A. S. Barr, Chairman, Professor of Education, University of Wisconsin, Madison 6, Wis.

Arthur T. Jersild, Professor of Education, Teachers College, Columbia University, New York City. Editorially responsible for materials on child welfare, guidance, and development, published each December.

Palmer O. Johnson, Professor of Education, University of Minnesota, Minneapolis, Minnesota. Editorially responsible for materials on measurements, statistics, and methods of experimental research, published each March.

H. H. Remmers, Professor of Educational Psychology, Director Division Educational Reference, Purdue University, Lafayette, Indiana. Editorially responsible for materials on learning, teaching and supervision, published each September.

J. Wayne Wrightstone, Assistant Director, Bureau of Reference, Research and Statistics, Board of Education of the City of New York, 110 Livingston Street, Brooklyn, New York. Editorially responsible for materials on curriculum construction, published each June.

CONTRIBUTING EDITORS

Emmett A. Betts, Research Professor and Director of Reading Clinic, Pennsylvania State College, State College, Pennsylvania.

Gilbert L. Betts, Supervisor of Graduate Research in Education, Colorado State College, Fort Collins, Colorado.

Leo J. Brueckner, Professor of Education, University of Minnesota, Minneapolis, Minnesota.

Oscar K. Buros, Associate Professor of Education, Rutgers University, New Brunswick, New Jersey.

Guy T. Buswell, Professor of Educational Psychology, University of Chicago, Chicago, Illinois.

Harold D. Carter, Associate Professor of Education, University of California, Berkeley 4, California.

Leslie L. Chisholm, Associate Professor of Education, State College of Washington, Pullman, Washington.

Herbert S. Conrad, Technical Consultant, College Entrance Examination Board, Princeton, New Jersey.

Stephen M. Corey, Professor of Educational Psychology, University of Chicago, Chicago, Illinois.

Edward E. Cureton, Secretary-Treasurer, Richardson, Bellows, Henry & Company, Inc., New York 4, New York.

Robert A. Davis, Professor of Education, Director of Bureau of Educational Research, University of Colorado, Boulder, Colorado.

Harl R. Douglass, Director of College of Education, University of Colorado, Boulder, Colorado.

Harold A. Edgerton, Director, Occupational Opportunities Service, Professor of Psychology, Ohio State University, Columbus 10, Ohio.

John C. Flanagan, Professor of Psychology, University of Pittsburgh, Pittsburgh, Pennsylvania.

Carter V. Good, Dean, Teachers College, University of Cincinnati, Cincinnati 21, Ohio.

Robert W. B. Jackson, Assistant Professor of Educational Research; Assistant Director, Department of Educational Research, Ontario College of Education, University of Toronto, Toronto, Canada.

Harold E. Jones, Professor of Psychology and Director, Institute of Child Welfare, University of California, Berkeley 4, California.

Noel Keys, Professor of Education and Lecturer in Human Relations, University of California, Berkeley, California.

D. Welty Lefever, Professor of Education, University of Southern California, Los Angeles, California.

Edward A. Lincoln, Consulting Psychologist, Halifax, Massachusetts.

Irving Lorge, Professor of Education, Executive Officer, Institute of Psychological Research, Teachers College, Columbia University, New York 27, New York.

A. R. Mead, Director of Educational Research, University of Florida, 330 P. K. Yonge Building, Gainesville, Florida.

T. E. Newland, Lt. Comdr., USNR, 2702 Wisconsin Avenue, N. W., Washington 7, D. C.

C. W. Odell, Professor of Education, University of Illinois, Urbana, Illinois.

Willard C. Olson, Professor of Education, Director of Research in Child Development, University of Michigan, Ann Arbor, Michigan.

W. E. Peik, Dean and Professor of Education, University of Minnesota, Minneapolis, Minnesota.

Walworth R. Plumb, Director Student Personnel Service, University of Minnesota (Branch), Duluth, Minnesota.

S. L. Pressey, Professor of Educational Psychology, Ohio State University, Columbus, Ohio.

Clarence E. Ragsdale, Professor of Education, University of Wisconsin, Madison, Wisconsin.

William Reitz, Associate Professor of Education, College of Education Examiner, Wayne University, Detroit 1, Michigan.

Henry D. Rinsland, Professor of Education and Director of Educational Research, The University of Oklahoma, Norman, Oklahoma.

Robert T. Rock, Jr., Professor of Psychology, Head of Dept. of Psychology, Graduate School, Fordham University, New York City.

Phillip J. Rulon, Professor of Education, Harvard Graduate School of Education, Cambridge 38, Mass.

Douglas E. Scates, Professor of Education, Duke University, Durham, North Carolina.

Harold Seashore, Director, Test Division, The Psychological Corporation, New York 18, New York.

David Segal, Educational Consultant, Specialist in Tests and Measurements, Federal Security Agency, U. S. Office of Education, Washington, D. C.

Paul W. Terry, Professor of Educational Psychology, University of Alabama, University, Alabama.

Helen Thompson, Associate Attending Psychologist, New York Post-Graduate Hospital, 303 East 20th Street, New York 3, New York.

Robert L. Thorndike, Associate Professor of Education, Teachers College, Columbia University, New York City.

Herbert A. Toops, Professor of Psychology, Ohio State University, Columbus, Ohio.

Maurice E. Troyer, Director, Bureau of School Services, Syracuse University, Syracuse 10, New York.

Helen M. Walker, Professor of Education, Teachers College, Columbia University, New York City.

Beth L. Wellman, Professor of Psychology, Child Welfare Research Station, State University of Iowa, Iowa City, Iowa.

Guy M. Wilson, Professor of Education, Emeritus, Boston University, Boston, Massachusetts.

Paul A. Witty, Professor of Education, Director of Psycho-Educational Clinic, School of Education, Northwestern University, Evanston, Illinois.

Ernest R. Wood, Professor of Education, New York University, New York City.

D. A. Worcester, Chairman Department of Educational Psychology and Measurement, University of Nebraska, Lincoln, Nebraska.

LITHO PRODUCTIONS

MADISON WIS

Journal of Experimental Education

Volume XVIII September, 1949 Number 1

A FOLLOW-UP OF SOME EARLIER WISCONSIN STUDIES OF TEACHING ABILITY

WILLARD JOSEPH BRANDT*
University of Wisconsin

Section I

THE PROBLEM

Introduction

The problem of measurement and prediction of teaching efficiency has become increasingly important since the turn of the century. During the period from 1930 to 1940, the teacher training institutions were graduating about 50,000 more teachers than were needed. Some of the teacher training institutions were attempting to select the students entering the teaching profession. School superintendents were selecting the individuals whom they thought would do the best job. Most of the selections were made on the basis of subjective data.

Almost immediately following the period during which there was a large surplus of qualified teachers came the war period with a shortage of teachers. Many left the teaching profession for more lucrative positions or were inducted into the armed services. The questions quite generally asked were: Are the best teachers leaving the profession? Did the teaching profession improve during the war period by losing those who did not enjoy teaching and who may not have been the most successful? How may better teachers be recruited and educated for teaching?

It has been estimated by the National Education Association that about 100,000 fewer elementary teachers than needed have been trained each year for the past two years. This shortage may not, however, continue for many years. To meet this present shortage of elementary teachers, the teacher training institutions have organized recruitment programs. What criteria should they use in selecting prospective teachers?

This follow-up investigation is concerned with the validity and reliability of predictions of teaching success, made by C. V. LaDuke[1], L. E. Rostker[2], R. D. Jones[3], and L. J. Lins[4].

* The author wishes to express his deep appreciation to his major professor, Dr. A. S. Barr, for his introduction to the problem and for his advice, aid and encouragement when progress was slow. He was most generous with his time and guidance. Appreciation is extended to C. V. LaDuke, L. E. Rostker, R. D. Jones, and L. J. Lins for their assistance in obtaining data to conduct this study. The author is also indebted to the teachers, principals, supervisors, graduate students, and the many others who assisted in this follow-up study.

1. C. V. LaDuke, "The Measurement of Teaching Ability," Journal of Experimental Education, XIV (September 1945), pp. 75-100.

2. L. E. Rostker, "The Measurement of Teaching Ability," Journal of Experimental Education, XIV (September 1945), pp. 6-51.

3. R. D. Jones, "Prediction of Teaching Efficiency from Objective Measures," Journal of Experimental Education, XV (September 1946), pp. 85-99.

4. L. J. Lins, "The Prediction of Teaching Efficiency," Journal of Experimental Education, XV (September 1946), pp. 2-60.

Have important changes occurred during the interim between the time when these investigations were conducted and the current status of the teachers involved in these studies?

Statement of the Problem

The four studies, with which this follow-up study is concerned, had in general the same objectives. They differed, however, in the time at which each study was conducted. In a sense, the LaDuke[5] and Rostker[6] studies constituted the background for the Jones[7] and Lins[8] studies. The LaDuke and Rostker studies attempted to discover the interrelationships of certain factors--after the teachers were in the field, which might be employed in future prediction studies. The Jones and Lins studies started with the people during the pre-service period and attempted to ascertain the validity and reliability of measures commonly used to predict teaching success in the field. All of these studies attempted to find the interrelationships between measures when applied within short intervals of time. The follow-up study attempts to find the interrelationships between measures after a lapse of eight years for the LaDuke study, eleven years for the Rostker study, and four years for the Lins and Jones studies.

The purposes of this follow-up study then are: (1) to find the interrelationships at the present time between some of the measures employed in the original studies, (2) to determine if there have been any significant changes in the relationships found between certain measures of teaching ability and efficiency as reported in the original studies, after a lapse of time, and (3) if the predictions of teaching success made in the original studies in terms of the measures employed are valid and reliable.

A secondary purpose of this follow-up study is to ascertain whether there is any relationship between teaching success as measured by the criterion of this follow-up study and six qualitative factors obtained from a Data Booklet (Appendix B1).*

Section II

GENERAL DESIGN OF THE STUDY

Selection of Participating Teachers

The teachers participating in this follow-up study on the measurement and prediction of teaching efficiency were chosen from those who had participated in the previous studies made by LaDuke[9], Rostker[10], Jones[11], and Lins[12], and who had continued in teaching until May, 1947. Of the 185 teachers who participated in the original studies, 80 were still teaching and indicated a willingness to cooperate in another study.

A rating on the Wisconsin Adaptation of the M-Blank, was obtained for 57 of the 80 teachers. Thirty-five of the 57 teachers, for whom the M-Blank rating had been secured, completed and returned the Data Booklets.

The information on the 35 teachers who returned the Data Booklets is analyzed separately.

Collection of Teacher Data

The first correspondence with the participants was a postal reply card (Appendix B2), which asked for information so that the files in the Teacher Personnel Research Bureau might be brought up-to-date. If no reply was received to the first card, it was followed by a letter and another card which asked for the same information.

A follow-up card, sent to all teachers who had returned the first information card, asked them if they were willing to cooperate in this study. Of the 185 who had participated in one or another of the four previous investigations, 80 expressed

5. LaDuke, op. cit.

6. Rostker, op. cit.

7. Jones, op. cit.

8. Lins, op. cit.

* All references to Appendices may be found in original thesis on file in University of Wisconsin Library.

9. LaDuke, Ibid.

10. Rostker, Ibid.

11. Jones, Ibid.

12. Lins, Ibid.

a willingness to cooperate in another study. Of those willing to participate in this study, nine were of the original 34 in the LaDuke study, 17 of the 28 in the Rostker study, 35 of the 65 in the Jones study, and 19 of the 58 in the Lins study.

In cooperation with Miss Frances Kline and other graduate students enrolled in the School of Education at the University of Wisconsin during 1946-47, the original Student Data Booklet[13] and the Supplementary Data Booklet[14] used by Lins and Jones were revised and combined into a new Data Booklet (Appendix B1).

Each teacher who had indicated a willingness to cooperate in this study was sent a copy of the new Data Booklet and asked to supply the information and return it as soon as possible. Of the 39 Data Booklets returned, four were eliminated because they were incomplete. The information in the Data Booklet included such items as age, sex, years of teaching experience, training beyond high school, type of teaching certificate, beginning teaching salary, salary in 1947, parental occupation, work experience other than teaching, participation in music, dramatics, community activities, reading interests and membership in professional organizations (Appendix B1).

Criterion and Other Measures Applied to the Teachers in the Follow-Up Study

The criterion of teaching efficiency used in this follow-up study is a supervisory rating made on the Wisconsin Adaptation of the M-Blank (Appendix B5). The M-Blank was also used as a measurement of teaching efficiency in the original studies made by Lins and Jones.

A letter of explanation and directions was sent with a copy of the M-Blank to the principal, superintendent or supervisor who was responsible for the supervision of the individual participating in this study. Fifty-seven of the 80 M-Blanks sent to the supervisors were completed and returned.

The final evaluations on the M-Blank were made on a five-point scale of five for outstanding, four for above average, three for average, two for below average and one for poor or inferior.

Additional ratings on the participants were obtained from the offices of the Teacher Placement Department of the Wisconsin Industrial Commission, the Teacher Placement Bureau of the University of Wisconsin, and the State Department of Public Instruction.

The evaluation sheet used by the Industrial Commission (Appendix B7), employed a scale of five categories: poor, below average, average, above average, and outstanding. The Placement Bureau of the University of Wisconsin (Appendix B8) used a scale of four categories: poor, fair, good, and excellent, with subdivisions under each category. These subdivisions were arbitrarily divided by the writer into five parts to correspond with the other scales. The State Department of Public Instruction used the letters A, B, C, D, and E to rate the efficiency of the teachers. Since the three rating sheets used by the Industrial Commission, the University of Wisconsin, and the State Department of Public Instruction were very similar, the five-point scale used in the M-Blank was used with these.

The evaluations made on the rating sheets used by the Industrial Commission, the University of Wisconsin, and the State Department of Public Instruction were obtained from their records. These ratings had been made by previous principals or supervisors, with the exception of the State Department of Public Instruction, whose supervisors had made the ratings.

Ratings for 46 participants were secured on the Industrial Commission's rating sheet, for 23 teachers on the blank used by the Placement Bureau of the University of Wisconsin, and for 27 teachers by the State Department of Public Instruction.

The predictions of teaching efficiency obtained, when using the same regression equation and independent variables employed in the original studies, will be compared with the actual scores received on the M-Blank in this follow-up study.

The follow-up study also attempts to ascertain if there are any significant relationships between the criterion and the following qualitative traits:

1. The father was or was not a farmer.
2. The individual did or did not participate in athletics in high school and/or college.
3. The individual did or did not participate in musical activities in high school and/or college.

13. L. J. Lins, Prediction of Teaching Efficiency, Unpublished Thesis, University of Wisconsin Library, Madison, Wisconsin, 1946.

14. R. D. Jones, The Prediction of Teaching Efficiency from Objective Measures, Unpublished thesis, University of Wisconsin Library, Madison, Wisconsin, 1946.

4. The individual did or did not participate in dramatics in high school and/or college.

5. The individual did or did not participate in community activities as a teacher.

6. The individual did or did not belong to a professional organization as a teacher.

When analyzing the above named six qualitative factors, the information for the teachers in the original studies by LaDuke and Rostker was combined, and also the information relative to the Jones and Lins studies was combined.

Although elsewhere in this study comparisons made between this follow-up study and the four preceding investigations have been separate, it seemed better and satisfactory to combine them here.

Section III

ANALYSIS OF THE DATA

Study Number One - LaDuke

Teacher Measures

The criterion of supervisory ratings used in this follow-up study was made on the Wisconsin Adaptation of the M-Blank. In addition to the M-Blank ratings, ratings for three teachers were secured on the evaluation sheet used by the Teacher Placement Department of the Wisconsin Industrial Commission.

The correlations reported by LaDuke[15] between the supervisory rating scales used in the original study and the criterion of pupil gain employed in the original study when the rating was made by the county superintendent, were as follows:

Criterion with Almy-Sorenson Rating Scale	-.01
Criterion with Michigan Rating Scale	.10
Criterion with Torgerson Rating Scale	.00

When the ratings of the teachers were made by supervising teachers, the correlations between the criterion of pupil gain employed in the original study and the supervisory ratings used in the original study were:

Criterion with Almy-Sorenson Rating Scale	-.24
Criterion with Michigan Rating Scale	-.30
Criterion with Torgerson Rating Scale	-.14

The zero order correlation coefficient between the criterion of pupil gain employed in the original study and the criterion of supervisory ratings employed in this follow-up study, however, is .14 . Although this correlation between the criterion of pupil gain in the original study and the criterion of supervisory ratings in this follow-up study is not large enough to be statistically significant, it does seem to indicate a closer relationship between the measures after a lapse of eight years. It appears that the predictions of teaching success made by LaDuke have improved rather than deteriorated over a period of years.

The correlations between the supervisory ratings made by county superintendents in the original study and the supervisory ratings in this follow-up study, Table I, were:

M-Blank with Almy-Sorenson Rating Scale	.81
M-Blank with Michigan Rating Scale	.77
M-Blank with Torgerson Rating Scale	.73

Using the ratings made by the supervising teachers, the correlations were:

M-Blank with Almy-Sorenson Rating Scale	.42
M-Blank with Michigan Rating Scale	.46
M-Blank with Torgerson Rating Scale	.71

Although the correlations between the supervisory ratings of the original and this follow-up study are high, they are not statistically significant.

A low correlation between pupil gain and supervisory ratings also agrees with the find-

15. C. V. LaDuke, "Measurement of Teaching Ability," Journal of Experimental Education, XIV (September 1945), p. 97.

TABLE I

ZERO COEFFICIENTS OF CORRELATION, STANDARD ERROR OF CORRELATION AND THE LEVEL OF SIGNIFICANCE BETWEEN THREE RATING SCALES EMPLOYED BY LADUKE AND THE M-BLANK RATING USED IN THE FOLLOW-UP STUDY

	N	Ratings Made by County Superintendents			Ratings Made by County Supervisors		
		r	Er	Percent Level of Significance*	r	Er	Percent Level of Significance*
Michigan Rating Scale	7	.77	± .18	7.02	.42	± .37	35.24
Torgerson Rating Scale	7	.73	± .21	8.50	.46	± .35	30.30
Almy-Sorenson Scale	7	.81	± .16	4.98	.71	± .22	11.42

*For the purpose of consistency, the formula used by Linn to obtain the percent level of significance, was also employed in this follow-up study. However, if the formula in C. C. Peters and W. R. VanVoorhis, Statistical Procedures and Their Mathematical Bases, page 154 (which is for small samples) is used the percent level of statistical significance is increased in all cases.

ings of Barr[16], who obtained coefficients of correlation ranging from .09 to .13 between supervisory ratings and pupil gain, and Gotham[17] found a correlation coefficient ranging from .35 to .43 using the same supervisory rating scales as LaDuke.

Data Booklet Information - LaDuke Study

A better understanding of the participants of this investigation can be obtained by studying some of the information obtained from the Data Booklet (Appendix B1).

Table II shows that the number of schools in which the participants had taught varied from four to nine, with a mean of six. About three years was the usual period of tenure of a teacher in any one school. The lowest beginning annual salary was $585, and the highest beginning annual salary was $900, with a mean beginning annual salary of $772.50. In 1947, the highest annual salary was $2475, and the lowest annual salary was $2025, with $2153.33 as the mean annual salary. It is interesting to note that the teacher who received the lowest beginning salary ($585) received the highest annual salary ($2475) in 1947, or an increase of over 425 percent, and the teacher who had received the highest beginning annual salary ($900) received the lowest annual salary ($2025) in 1947, or an increase of only 225 percent. It should also be noted in Table II that it was not the teacher who moved most frequently who received the highest salary in 1947.

In 1939, all of the participating teachers had county one-year licenses. In 1947, two had life certificates, two had five-year state licenses, and one had a state one-year license.[18] The two teachers with life certificates had completed three additional years of schooling.

The Pearson product-moment correlation coefficient between the criterion of supervisory rating used in this follow-up study and the salary received in 1947 is .59 $\pm$ .30, which is significant at the 19.02 percent level (Table XIX, Appendix A). Even though the correlation coefficient is not large enough to be statistically significant, it might be said that it substantiates a popular opinion that there is not too much relationship between teaching efficiency and salary received. It also agrees with the findings of other previous studies. F. B. Knight[19] found a correlation coefficient of .35 between salaries and teaching efficiency as rated by supervisors, mutual judgments of the teachers themselves and consensus of pupil opinions. R. L. Morton[20] obtained a correlation of coefficient of .24 between salaries and the criterion of pupil reactions in the cities studied.

Statistical treatment of some of the information in the Data Booklet, to determine if there are any significant relationships between the criterion and six selected qualitative variables, will be included in Study Number Two of this section.

Study Number Two - Rostker

Criterion and Other Teacher Measures

The criterion of teaching ability employed in this part of the follow-up study was that of supervisory ratings. Of the 28 teachers participating in the original study in 1936, 19 were still teaching in 1947, and supervisory ratings on the Wisconsin Adaptation of the M-Blank were obtained for 12 of the 19.

The correlation between the criterion of pupil gain used in the original study and the criterion of supervisory ratings used in the follow-up study is .35 $\pm$ .27. In the original study by Rostker[21] the coefficients of correlation between the criterion of pupil gain and the supervisory ratings, Table III, on three different scales were:

Pupil gain with Almy-Sorenson Rating Scale	-.02

16. A. S. Barr, "The Validity of Certain Instruments Employed in the Measurement of Teaching Ability," The Measurement of Teaching Efficiency (New York: Macmillan Co., 1935), pp. 71-141.

17. R. E. Gotham, "Personality and Teaching Efficiency," Journal of Experimental Education, XIV (December 1945), pp. 157-165.

18. Since 1939, a state one-year license is equivalent to a county one-year license.

19. F. B. Knight, Qualities Related to Success in Teaching, Contributions to Education, No. 120 (New York: Teachers College, Columbia University, 1922).

20. R. L. Morton, "Qualities of Merit in Secondary Teachers," Educational Administration and Supervision, V (May 1919), pp. 225-238.

21. Rostker, op. cit.

TABLE II

TEACHER INFORMATION - LADUKE STUDY (1937-1939)

Teacher No.	Sex	Age	Years of Teaching	Yrs. Training Beyond High School	Certificate	No. of Schools Taught in	Subject or Grade Taught	Beginning Salary	1947 Salary
51	F	41	21	5	L	8	Grades	$900	$2270
52	F	33	10	1	1 Yr.	4	Grades	765	2025
54	F	34	14	4	L	6	Grades	585	2475
55	F	33	15	1	5 Yrs.	4	Grades	675	2025
56	F	44	17	4	L	9	Grades	810	2100
57	F	42	25	1	5 Yrs.	6	Grades	900	2025
			M 17			M 6.1		M $772.50	$2153.33

TABLE III

ZERO ORDER COEFFICIENTS OF CORRELATION, STANDARD ERROR OF CORRELATION AND LEVEL OF SIGNIFICANCE BETWEEN THREE RATING SCALES EMPLOYED BY ROSTKER IN HIS ORIGINAL STUDY AND THE M-BLANK RATING EMPLOYED IN THE FOLLOW-UP STUDY

	N	r E_r	Percent Level of Significance
Michigan Rating Scale	12	$.27 \pm .28$	36.82
Torgerson Rating Scale	12	$.19 \pm .20$	53.52
Almy-Sorenson Rating Scale	12	$.11 \pm .27$	73.00

TABLE IV

COMPARISON OF PREDICTED SCORES USING SIX VARIABLES WITH THE ACTUAL M-BLANK SCORES IN THE FOLLOW-UP STUDY (ROSTKER)

Teacher Number	Predicted Score	Actual M-Blank Score	Difference
39	3.61	2.88	-.73
40	3.62	4.11	.49
41	3.60	3.14	-.46
42	3.32	4.22	.90
43	4.32	3.61	-.71
44	2.98	3.22	.24
45	3.61	3.28	-.33
46	3.91	4.11	.20
47	3.69	3.94	.25
48	4.27	4.72	.45
49	4.27	3.33	-.94
50	3.87	4.33	.46

Pupil gain with Michigan Rating Scale	.15
Pupil gain with Torgerson Rating Scale	-.01

The correlation coefficients between the supervisory rating scales used in the original study and the supervisory ratings of this follow-up study are:

M-Blank with Almy-Sorenson Rating Scale	.11
M-Blank with Michigan Rating Scale	.27
M-Blank with Torgerson Rating Scale	.19

Although the correlations in the above instances are not large enough to be statistically significant, the larger correlation of .35 between the supervisory ratings of this follow-up study and the criterion of pupil gain used in the original study shows that a closer relationship does seem to be present between the two measures, applied after a lapse of 11 years.

Employing the general form for the multiple regression equation which can be written as:

$$X_c = B_{c_1}\frac{\sigma c}{\sigma 1}X_1 + B_{c_2}\frac{\sigma c}{\sigma 2}X_2 \ldots + B_{c_n}\frac{\sigma c}{\sigma n}X_n$$

$$+ (M_c - B_{c1}\frac{\sigma c}{\sigma 1}M_1 - B_{c2}\frac{\sigma c}{\sigma 2}M_2 \ldots$$

$$- B_{c_n}\frac{\sigma c}{\sigma n}M_n)$$

in which X_c indicates the criterion or dependent variable and X_1, X_2, and X_n the independent variables. The proper calculations as indicated by the equation give the following multiple regression equation in raw score form:

$$X_0 = .008\,X_1 + .004\,X_2 + .001\,X_3 + .359\,X_4 + .017\,X_5 - .003\,X_6 - .89$$

The scores obtained by the individual teachers on the six teacher measures used, and reported by Rostker[22], were substituted for the independent variables in the regression equation. A comparison of the predicted scores and actual scores obtained by the teachers is found in Table IV. The actual scores deviate from the predicted scores within the range of -.94 to +.90. Eight of the scores deviate from the mean by less than 0.5 of a point, and they all deviate by less than 1.0. The mean deviation between the predicted score and the actual score in this part of the follow-up study is .51, which is about the same as found for the Jones and Lins studies, which are reported in Tables XI, XII, XVII, and XVIII respectively in the follow-up study. The increased relationship, cited previously, between the criterion of pupil gain in the original study and the criterion of this follow-up study, and the small mean deviation of the actual scores in this follow-up study from the predicted scores when using six teacher measures employed by Rostker, seems to indicate that the predictions of teaching ability have improved rather than deteriorated during the period between the two studies.

Data Booklet Information - Rostker Study

Information concerning the participants of the Rostker study[23], found in Table V, shows that the number of schools in which the teachers have held positions range from three to ten, with a mean of about four years tenure in any school. Their beginning annual teaching salary ranged from $225 to $1500. Much of this variance is due to the greater number of years some have been teaching. The number of years in teaching ranged from 19 to 42, with a mean of 27.5 years. The lowest annual salary in 1947 was $1710, while the highest was $2974. The mean beginning annual salary was $742.50, and the mean annual salary in 1947 was $2457.90. The individual who had taught 42 years had increased his salary about 250 percent. All of the participants had life certificates in 1947.

A Pearson Product-Moment coefficient of correlation between the amount of salary received in 1947 and the criterion was .85 ± .09, which is significant at the 00.78 percent level, (Table XIX, Appendix A). This seems to indicate that there may be some degree of relationship between teaching efficiency as determined by the criterion of this study and their annual salary. The significant correlation between teaching efficiency as measured by the criterion of this follow-up study and salary as found here, however, is not in agreement with the findings in the sections of this follow-up study which are concerned with the LaDuke and Jones studies. As has been previously stated, most investigations have shown an insignificant relationship between salary and teaching efficiency.

22. L. E. Rostker, Measurement of Teaching Ability, Unpublished Thesis, University of Wisconsin Library, Madison, Wisconsin, 1941.

23 L. E. Rostker, "Measurement of Teaching Ability," Journal of Experimental Education, XIV (September 1945), pp. 6-51.

TABLE V

TEACHER INFORMATION - ROSTKER STUDY (1936)

Teacher No.	Sex	Age	Years of Teaching	Yrs. Training Beyond High School	Certificate	No. of Schools Taught in	Subject or Grade Taught	Beginning Salary	1947 Salary
39	F	49	27	3	L	3	Elem. Grades	$1050	$1710
40	M	43	21	5	L	7	Elem. Grades	855	2750
41	F	52	31	4	L	9	Grades	495	1845
42	F	58	31	4	L	6	Grades	270	3300
43	F	49	32	4	L	8	Grades	440	2974
44	M	42	20	4	L	6	Grades	1500	2700
46	F	41	19	5	L	3	Grades	1080	2500
48	F	59	42	5	L	3	Grades	225	2600
49	M	65	34	4	L	10	Grades	810	2100
50	F	37	18	4	L	3	Grades	700	2100
						M = 5.8		M =$ 742.50	M= $2457.90

In this part of the follow-up study, as in the section concerning the LaDuke investigation, the data obtained does not warrnat postulations as to whether the teachers concerned have become more or less efficient in classroom instruction.

Six Qualitative Factors Combining the LaDuke and Rostker Studies

Part of the information from the Data Booklets which the investigator wanted to use could not be numerically expressed. A biserial coefficient of correlation was accordingly used to determine the amount of relationship between the criterion, numerically expressed, and the other variable expressed in two alternative categories[24]:

$$r_{bis} = \frac{\bar{M}_2 - \bar{m}_1}{\sigma} \frac{pq}{y}$$

$\bar{m}$ = mean of smaller frequency distribution

$\bar{M}$ = mean of larger frequency distribution

σ = standard deviation of total frequency distribution

where: p = proportion of the total frequency of the larger distribution to the grand total frequency.

q = 1 - p

y = height of the ordinate of a normal curve of unit area and unit standard deviation at the point separating the area of the curve into proportion p and q.

The processes observed in calculating biserial r are based on some of the properties of the normal probability curve. It was assumed that the dichotomous traits were normally distributed and continuous, and also that the relationship between the two factors was linear.

A biserial coefficient of correlation was calculated for the following qualitative traits:

1. The father was or was not a farmer.
2. The individual did or did not participate in athletics in high school and/or college.
3. The individual did or did not participate in musical activities in high school and/or college.
4. The individual did or did not participate in dramatics in high school and/or college.
5. The individual did or did not participate in community activities as a teacher.
6. The individual did or did not belong to a professional organization as a teacher.

The standard error of the biserial coefficients of correlation, the biserial correlation coefficient and the percent level of significance is found in Table VI.

An examination of Table VI shows a biserial coefficient of correlation of .35 ± .22 between the criterion of the follow-up study and whether the father was or was not a farmer. This is significant at the 26.16 percent level. Even though the coefficient of correlation is not large enough to be statistically significant at the one or five percent level, there seems to be some relationship between their success as rural and/or elementary teachers and the fact that their father was a farmer. A closer analysis of the data in the LaDuke and Rostker studies shows that all the participants were elementary teachers, and all would be called rural if the sociological definition of a rural community being one of less than 2500 population were used.

The biserial coefficient of correlation between the criterion of the follow-up study and participation in athletics in high school and/or college is .37 ± .22, and between participation in musical activities is .09 ± .26, neither of which is significant at the five percent level or less. This would seem to indicate that there is very little relationship between success in teaching and participation in athletics or musical activities in high school and/or college. A negative biserial coefficient of correlation, -.47 ± .20 is found between the criterion of this follow-up study and participation in community activities as a teacher, which is significant at the 7.04 percent level.

Seagoe[25] found a correlation of coefficient of -.05 between teaching success and musical talent, as determined by percentile rank made by the administrator. A zero order coefficient of correlation of .37 between personal data and

24. T. C. McCormick, Elementary Social Statistics (New York: McGraw-Hill Book Co., 1941), p. 201.

25. M. V. Seagoe, "Prediction of In-Service Success in Teaching," Journal of Educational Research, XXXIX (May 1946), pp. 658-663.

TABLE VI

BISERIAL COEFFICIENTS OF CORRELATION, STANDARD ERROR OF CORRELATION AND LEVEL OF SIGNIFICANCE BETWEEN SIX DICHOTOMIES AND THE CRITERION OF THE FOLLOW-UP STUDY (LADUKE AND ROSTKER STUDIES COMBINED)

Dichotomy	r E_r	Percent Level of Significance
Father was or was not a farmer	.36 ± .22	26.16
Did or did not participate in athletics in high school and/or college	.37 ± .22	15.56
Did or did not participate in musical activities in high school and/or college	.09 ± .2(	72.98
Did or did not participate in dramatics in high school and/or college	.09 ± .26	72.64
Did or did not participate in community activities as a teacher	-.47 ± .20	7.04
Did or did not belong to a professional organization	.23 ± .24	37.36

N = 16

TABLE VII

ZERO ORDER COEFFICIENTS OF CORRELATION, STANDARD ERROR OF THE CORRELATION, AND SIGNIFICANCE LEVEL OF THE CORRELATION BETWEEN THE INCREASE IN SALARY SINCE BEGINNING TEACHING AND THE NUMBER OF TEACHING POSITIONS HELD

	N	r E_r	Percent Level of Significance
Lins Study	7	.05 ± .41	91.24
Jones Study	12	.66 ± .17	3.00
Rostker and LaDuke Studies	16	.09 ± .27	73.38

musical talent was found by Durlinger[26]. Bossing[27] and Boyce[28] found correlation coefficients of .06 and .57 between supervisory rating and attitude and interest in the community. A coefficient of correlation of .41 between supervisory ratings and extra-curricular activities was found by Somers[29]. These results are not greatly different from those found in the follow-up study, which would seem to indicate that the relationship between the factors considered does not change significantly over a period of years.

A Pearson Product-Moment coefficient of correlation was also calculated to see if there was any significant relationship between the increase in salary since beginning teaching and the number of teaching positions held. The coefficient of correlation, found in Table VII, is .09 ± .27, which is significant at the 73.38 percent level.

Summary and Conclusions

As a result of this part of the follow-up study of the teachers who participated in the Rostker study, certain conclusions seem warranted:

1. Although the correlations between the measures employed in the original study and the measures employed in this follow-up study are not large enough to be statistically significant, the relationship has become closer than it was in the original study. Thus, it appears that the predictions of teaching efficiency made in the original studies have improved during the lapse of time rather than deteriorated.

2. The criterion used in the follow-up study did not show any significant change in the teaching efficiency of the individuals concerned since the previous studies were conducted in 1936 and 1939.

3. The zero order coefficient of correlation between the criterion of pupil gain in the original study by Rostker and the criterion of supervisory ratings used in this follow-up study was not significant.

4. No significant coefficients of correlation were found between the supervisory ratings obtained by Rostker and the criterion of this follow-up study.

5. For the LaDuke participants, there is no significant correlation between salary paid in 1947 and success as a teacher, as found with the M-Blank, while for the Rostker participants the correlation is highly significant.

6. There seems to be no relationship between the number of times a teacher changes positions and the salary increase.

7. There seems to be no significant relationship between the six dichotomous variables considered and success as measured by the criterion.

Study Number Three - Jones

Criterion and Other Teacher Measures

The primary purpose of this section of the follow-up study, as in the two preceding sections, is to determine if the predictions made by the original investigators were reliable and valid and to determine if the teachers included in this study made any significant changes in their teaching efficiency during the time which has lapsed between the first investigation and the present one.

The criterion of teaching efficiency in this follow-up study is supervisory ratings, which was also one of the criteria used by Jones in the original study.

As stated in Section II, the criterion of supervisory ratings were secured by sending copies of the Wisconsin Adaptation of the M-Blank to the principals of the schools where the teachers were located in 1947. A similar technique was employed by Jones.

Pearson Product-Moment coefficients of correlation between the criterion (M-Blank ratings) used in this follow-up study and one of the cri-

---

26. G. W. Durlinger, "Scholastic Prediction in a Teachers College," Journal of Experimental Education, XI (June 1943), pp. 257-267.

27. N. L. Bossing, Teacher Aptitude Tests and Teacher Selection, Research in Higher Education Bulletin, No. 12 (Washington, D. C.: U. S. Office of Education, 1932), pp. 117-133.

28. A. C. Boyce, "Qualities of Merit in Secondary School Teachers," Journal of Educational Psychology, III (March 1912), pp. 144-157.

29. G. T. Somers, Pedagogical Prognosis: Predicting the Success of Prospective Teachers, Contributions to Education, No. 140 (New York: Teachers College, Columbia University, 1923).

teria employed by Jones, namely, that of supervisory ratings, were calculated. The Wisconsin Adaptation of the M-Blank was also used by Jones to secure supervisory ratings.

The coefficient of correlation between the supervisory ratings obtained on the M-Blank by Jones and the supervisory ratings obtained on the M-Blank in this follow-up study, as shown in Table VIII, is .48 ± .17, which is significant at the 2.66 percent level. This tends to substantiate further Lins' findings regarding the reliability and validity of the M-Blank which, as stated in Section II, Lins found significant. Whatever the M-Blank measures, it seems to do so with a high degree of reliability and after a period of years.

A comparison of the M-Blank ratings obtained in this follow-up study and the M-Blank ratings secured by Jones is found in Table IX. It shows a difference of .11 in the means of the two ratings. Using the equation for the critical ratio between the means:[30]

$$C.R. = \frac{M_1 - M_2}{E^2_{m_1} + E^2_{m_2} - 2r_{12} E_1 E_2}$$

it is found that the difference is significant at the 80.08 percent level. Accepting the one or five percent level of significance as being satisfactory, the change in the teachers as measured by the M-Blank is not significant. However, Table IX shows that 16 of the teachers received a higher rating, with only six receiving a lower rating. It should also be noted that three teachers received a rating of more than one point less than the rating received in the original study using the five-point scale, explained in Section II, and one teacher increased her rating more than one point.

Although the mean difference between the ratings observed in the original study and ratings obtained in this follow-up study is not large enough to be significant, it does show a tendency for most of the teachers to be rated higher now after a period of years. Assuming that this gain is not due to the "halo effect," it does seem to indicate improvement in teaching ability.

Coefficients of correlation were also calculated between the ratings secured on the evaluation sheets used by the Teacher Placement Department of the Wisconsin Industrial Commission (Appendix B7); the Placement Bureau of the University of Wisconsin (Appendix B8); the State Department of Public Instruction (Appendix B6); and the criterion of this follow-up study.

The coefficient of correlation between the Industrial Commission blank and the criterion of this follow-up study, as shown in Table VIII, is .53 ± .19, which is significant at the 4.00 percent level. This would indicate that in only four samples in 100 might one expect to get a coefficient of this size by chance. The two rating sheets in this part of the study seem to be measuring to a degree the same qualities of teaching. Although the coefficients of correlation between the criterion of this follow-up study and the Rating Scale of the University of Wisconsin and the State Department of Public Instruction are .45 ± .21 and .65 ± .21, respectively, they are not significant. This lack of significant relationship with the criteria may be because different individuals made the ratings or that the rating sheets measure teaching qualities other than those measured by the M-Blank.

Data Booklet Information - Jones Study

Additional qualitative and quantitative data for 12 of the 22 teachers participating in this section of the study was obtained from the Data Booklet (Appendix B1). An observation of Table X, which has some of the data summarized, shows that the participants had taught a mean of 5.6 years, with a range varying from four to 11 years. Teacher Number 38 had taught previous to obtaining her degree from the University of Wisconsin in 1942. During the same period of years, the number of schools taught in was 2.5. The range was one to seven. The lowest annual beginning salary was $1000, and the highest beginning annual salary was $1800, with a mean beginning annual salary of $1269.58. In 1947, the lowest annual salary was $$1920, and the highest annual salary was $2920, with $2356.66 as the mean salary; thus, the mean gain in salary from 1942 to 1947 was nearly $1100.

The Pearson Product-Moment coefficient of correlation between the number of teaching positions held and increase in salary is .66 ± .17, and is significant at the 3.00 percent level, Table VII. A correlation of this size would be obtained by chance not more than three times in a hundred, which leads to the conclusion that the teacher who changes positions most frequently will increase her salary the most. It does not mean, however, that she will get the highest salary because she may have started at a much lower figure.

The coefficient of correlation between the

30. C. C. Peters and W. R. VanVoorhis, Statistical Procedures and Their Mathematical Bases (New York: McGraw-Hill Book Co., 1940), pp. 164-167.

TABLE VIII

ZERO ORDER COEFFICIENTS OF CORRELATION, STANDARD ERROR OF THE CORRELATION AND SIGNIFICANCE LEVEL OF THE CORRELATION BETWEEN THE CRITERION EMPLOYED IN THIS FOLLOW-UP STUDY AND THE OTHER RATING SCALES USED (JONES STUDY)

	N	r E_r	Percent Level of Significance
Jones M-Blank Rating	22	.48 ± .17	2.66
Industrial Commission Scale	16	.53 ± .19	4.04
University of Wisconsin Rating Scale	15	.45 ± .21	9.50
State Department of Public Instruction	9	.65 ± .21	6.58

salary received in 1947 and the criterion of this follow-up study is .38 ± .26, which is significant at the 20.78 percent level (Table XIX, Appendix A). Such a low correlation indicates that the best teachers, as rated by the criterion, do not necessarily receive the largest salary. It would also seem to indicate that the criterion measures different teaching qualities from those considered by the superintendent when granting salary increases. The low correlation between salary and teaching efficiency found here is in agreement with the findings of Knight[31] and Morton[32].

Statistical treatment of six dichotomous variables obtained from the Data Booklet will be found in Study Number Four of this section.

Prediction of Teaching Success - Jones Study[33]

The purpose of this section of the follow-up study differs in part from the other sections of the study previously discussed. The primary purpose of this section of the study is to ascertain if the predictions of teaching success made by Jones in the original study were valid and reliable.

From the several factors that were used by Jones in the original study for prediction of teaching success, three were selected for use in this follow-up study; namely, (1) freshman-sophomore grade point average, (2) junior-senior grade point average, and (3) the Bell Adjustment Inventory. The multiple regression equation using the above named factors gave a multiple correlation of .45, which, however, was not large enough to be statistically significant. The three above-listed variables were selected for use in this follow-up study because of the availability of the data needed.

The general form for the multiple regression equation was employed and the proper calculations as indicated by the equation give the following multiple regression equation in raw score form:

$$X_0 = .303\,X_1 + .417\,X_2 - .004\,X_3 + 2.28$$

The scores obtained by the teachers, as found by Jones, for the three variables that were employed for prediction in this study, were substituted for the independent variables in the multiple regression equation.

The predicted scores obtained, with the actual scores obtained in this follow-up study, are found in Table XI. This shows a range in the deviations of the actual score of -1.04 to .86, with a mean deviation--disregarding signs when adding--of .48.

A comparison of the predicted scores, with the actual score obtained by Jones in the original study, is found in Table XII. This shows

31. Knight, op. cit.

32. Morton, op. cit.

33. Jones, op. cit.

TABLE IX

A COMPARISON OF M-BLANK RATINGS IN JONES'S ORIGINAL STUDY AND THE M-BLANK RATINGS IN THIS FOLLOW-UP STUDY

Teacher Number	M-Blank Rating Follow-up Study	M-Blank Rating Jones Study	Difference
13	3.39	3.00	.39
14	4.27	3.75	.52
15	4.25	3.63	.62
16	4.61	4.00	.61
17	4.33	4.13	.20
18	4.44	4.00	.44
19	4.33	4.00	.33
20	4.55	4.13	.32
21	4.00	4.00	.00
22	3.94	4.38	-.44
23	3.72	4.75	-1.03
24	3.66	4.75	-1.09
25	2.88	3.94	-1.06
26	4.39	4.13	.26
27	3.00	2.25	.75
28	3.77	4.00	-.23
29	4.28	4.13	.15
30	3.61	3.00	.61
31	4.22	5.00	-.78
32	3.33	2.00	1.33
33	3.67	3.50	.17
34	4.11	3.75	.36
	M = 3.94 σ = .47	M = 3.83 σ = .23	

TABLE X

TEACHER INFORMATION - JONES STUDY (1943)

Teacher No.	Sex	Age	Years of Training	Yrs. Training Beyond High School	Certificate	No. of Schools Taught in	Subject or Grade Taught	Beginning Salary	1947 Salary
13	F	28	4	4	L	2	Science	$1080	$1920
14	F	27	4	4	L	2	Art	1350	2700
15	F	28	6	4	L	3	English	1080	2250
20	F	31	6	5	L	4		1000	2350
21	F	28	6	4	L	3	English	1080	2300
23	F	26	4	4	L	3	English	1400	2100
24	F	26	5	4	L	1		1665	2200
26	M	27	5	4	L	2	Agr. Biology	1800	2920
28	F	27	6	4	L	3	Foreign Language	1100	2300
34	F	27	5	4	L	3	History	1260	2250
36	F	30	6	4	L	2	Art	1050	2440
38	F	34	11	4	L	3 2	Grades Phy. Ed.	585* 1370	765* 2550
			M = 5.6			M = 2.5		$1269.58	$2356.66

* Not used in computing the averages.

TABLE XI

COMPARISON OF THE PREDICTED SCORES OF TEACHING EFFICIENCY USING THREE VARIABLES* AND THE ACTUAL SCORES OF TEACHING SUCCESS AS MEASURED BY THE M-BLANK IN THIS FOLLOW-UP STUDY (JONES)

Teacher Number	Predicted Score	Actual Score in Follow-Up	Difference
13	3.93	3.39	-.54
14	3.41	4.27	.86
15	3.90	4.25	.35
16	3.86	4.61	.75
18	3.76	4.44	.68
19	4.13	4.33	.20
20	4.13	4.55	.42
24	3.67	3.66	-.01
26	3.62	4.39	.77
27	4.04	3.00	-1.04
30	3.64	3.61	-.03
33	3.51	3.67	.16
			M = .48

*Three Variables:

1. Freshman-Sophomore Grade Point Average
2. Junior-Senior Grade Point Average
3. Bell Adjustment Inventory

a range in the deviations of the actual scores from the predicted scores of -1.79 to .11. The mean deviation between predicted and actual scores in the original study--disregarding signs when adding--is .51. The smaller mean deviation between the predicted and actual scores in this follow-up study than in the original study indicates a general improvement in the predictions after a period of years. Also, using the same rating scale in the original and this follow-up study to obtain the supervisory ratings, the actual scores of all except one of the teachers considered has improved.

In view of these results, it would seem that prediction of teaching success, using the three variables employed in the original and this follow-up study, is quite reliable after a period of years, and there also appears to be a tendency for the predictions of teaching efficiency to improve after a period of years.

Summary and Conclusion

This section of the follow-up study was primarily concerned with the reliability of the predictions of teaching efficiency as made in the original study and any significant change in teaching efficiency that has occurred during the lapse of time between the two studies. A secondary problem was to ascertain of certain selected variables, which are listed in Section II, have any relationship with teaching success.

The following conclusions seem to be warranted from the data obtained in the investigation:

1. The predictions of teaching efficiency made in the original study have improved rather than deteriorated during the lapse of years between the original study and this follow-up study.

2. This study has not found a significant change in the efficiency of teaching during the years between the two investigations. However, there is a tendency for the teachers to receive a higher rating in this follow-up study than they did in the original study.

3. There is a significant relationship between the M-Blank rating obtained in the Jones study and the M-Blank rating obtained in this follow-up study.

4. There is a significant relationship between the M-Blank rating and the rating on the Industrial Commission blank.

5. A significant relationship was not found between the criterion of this follow-up study and the ratings made on the blank employed by the University of Wisconsin and the Wisconsin Industrial Commission.

6. A significant relationship was found in this section between the number of teaching positions held and salary increases.

Study Number Four - Lins

Teacher Measures

As in the three preceding sections of this follow-up study, the primary purpose of this section is to determine if the predictions of teaching success made by Lins in the original study have improved or deteriorated during the years between the two studies and also to determine if there has been a significant change in teaching efficiency during the same period of time. A secondary purpose is to determine if there is any relationship between teaching success and the six variables mentioned previously in Section II.

The criterion for this section of the follow-up study is also that of supervisory ratings on the Wisconsin Adaptation of the M-Blank. As in Jones study, discussed in the previous section, supervisory ratings employing the M-Blank was one of the criteria used in the investigation made by Lins. The criteria of supervisory ratings employed by Lins, however, differed from the criteria of supervisory ratings of the Jones study and of this follow-up study in that it was a composite of ratings made on the M-Blank by five different raters.

The Pearson Product-Moment coefficient of correlation between the composite M-Blank rating of the Lins study and the M-Blank rating of this follow-up study is .73 $\pm$ .14, which is significant at the 1.64 percent level. However, a review of the Lins study shows that a coefficient of correlation of only .20 $\pm$ .17 was obtained between the composite rating on the M-Blank in 1943 and the ratings made by the principals in 1945.

Although there is a highly significant degree of relationship between the M-Blank ratings in the original study and in this follow-up study, the correlation between the M-Blank ratings obtained in the original study and the criterion of pupil gain was only .19. This low correlation is consistent with the results in the other three studies of this follow-up study, as well as with the findings of other studies previously cited.

A comparison of the composite M-Blank ratings obtained by Lins and the M-Blank criterion ratings obtained for this follow-up study is found in Table XIII. A difference of .24 be-

TABLE XII

COMPARISON OF THE PREDICTED SCORES OF TEACHING EFFICIENCY USING THREE VARIABLES* AND THE ACTUAL SCORES ON THE M-BLANK IN THE ORIGINAL STUDY BY JONES

Teacher Number	Predicted Score	Actual Score in Jones' Study	Difference
13	3.93	3.00	-.93
14	3.41	3.75	.34
15	3.90	3.63	-.27
16	3.86	4.00	.14
18	3.76	4.00	.24
19	4.13	4.00	-.13
20	4.13	4.13	.00
24	3.67	4.75	1.08
26	3.62	4.13	.51
27	4.04	2.25	-1.79
30	3.64	3.00	-.64
33	3.51	3.50	-.01
			M = .51

*Three Variables:

1. Freshman-Sophomore Grade Point Average
2. Junior-Senior Grade Point Average
3. Bell Adjustment Inventory

TABLE XIII

COMPARISON OF THE TEACHER RATINGS ON THE M-BLANK IN THE LINS STUDY AND THE M-BLANK RATINGS IN THE PRESENT FOLLOW-UP STUDY

Teacher Number	Present M-Blank Rating	Lins M-Blank Rating	Difference Gain of Present
1	4.11	4.25	-.13
2	2.94	2.75	.19
3	3.89	3.50	.39
4	3.39	3.25	.14
5	3.94	4.10	-.16
6	3.94	3.40	.54
7	4.26	4.10	.18
8	3.89	4.10	-.21
9	3.56	3.75	-.19
10	3.44	3.00	.44
11	4.33	4.00	.33
12	3.33	1.88	1.45
	M = 3.75	M = 3.51	.24
	σ = .42	σ = .67	

TABLE XIV

ZERO ORDER COEFFICIENTS OF CORRELATION, STANDARD ERROR, AND SIGNIFICANCE LEVEL BETWEEN TWO RATINGS USED IN THE FOLLOW-UP STUDY AND THE M-BLANK (LINS STUDY)

Rating Scale	r E_r	Percent Level of Significance
M-Blank Rating in the Lins Study	.72 $\pm$ .14	1.64
Rating Blank Used by the Wisconsin Industrial Commission	-.02 $\pm$ .45	97.90

tween the mean of the ratings in the original study and the mean of the follow-up study is found. Using the formula for the critical ratio between the means, it is found that the difference between the means is significant at the 61.02 percent level. The change as measured by the criterion is too small to be statistically significant; however, Table XIII shows that eight teachers have shown some improvement as measured by the criterion, and one of them increased nearly one and one-half points, using the five-point scale explained in Section II. Though the data available in this follow-up study is not sufficient to say definitely that the teachers have improved since the original study, the criterion employed does seem to indicate a general tendency toward improvement.

The coefficient of correlation between the rating blank used by the Teacher Placement Department of the Wisconsin Industrial Commission and the criterion is $-.02 \pm .45$, Table XIV. This correlation is too small to be significant in this case, and it also disagrees with the results obtained in the Jones study, in which a positive coefficient of correlation was obtained between the criterion of this follow-up study and the rating on the blank used by the Wisconsin Industrial Commission.

The low correlation between the Industrial Commission's blank and the criterion of this follow-up study, in contrast to the significant correlation between the previous composite M-Blank rating and the criterion of this follow-up study, may be due to the small number of cases and/or that different people made the ratings.

Data Booklet Information

The Data Booklet was constructed to obtain information which would help measure and predict teaching success. Some of the information obtained from the Data Booklet relative to the Lins study is found in Table XV. It shows a mean of four years of teaching, with a mean of 1.7 years in one teaching position. The number of teaching positions held varies from one to four. One teacher taught four years and in four different schools. The mean annual beginning salary was $1390, with $1230 as the lowest annual beginning salary and $1500 as the highest beginning salary. The mean salary received in 1947 was $2114.29, with $1875 as the lowest annual salary and $2350 the highest annual salary. The teacher who had moved four times in four years received the highest beginning annual salary and also the highest annual salary in 1947.

The Pearson Product-Moment coefficient of correlation between the amount of increase in salary, Table VII, since beginning teaching and the number of teaching positions held is $.05 \pm .41$, which is significant at the 92.24 percent level. From this we might conclude that the teachers who move frequently do not increase their salary more than the teachers who do not change teaching positions.

In an attempt to provide some evidence as to whether the most efficient teacher, as rated by the criterion, was receiving the highest salary, coefficients of correlation were calculated. The coefficients of correlation between the criterion and salary in 1947 is $-.70 \pm .20$, which is significant at the 00.86 percent level, (Table XIX, Appendix A). It appears that the better teachers among this group, as measured by the criterion, are receiving less salary than the poorer teachers. As has been pointed out in Study Number One of Section III, previous investigations have found a very low correlation coefficient between teaching efficiency and salaries.

Insufficient data prevented the calculation of coefficients of correlation for this part of the study between the criterion and the blanks used by the Placement Bureau of the University of Wisconsin and the State Department of Public Instruction.

Additional information obtained from the Data Booklet will be combined with the Jones study and analyzed in the following section.

Six Qualitative Factors Combining the Jones and Lins Studies

The information obtained from the Data Booklet and used here could not be numerically expressed. A biserial coefficient of correlation was accordingly employed in determining the amount of relationship between the criterion of this follow-up study, numerically expressed, and the other variable expressed in two alternative categories. As stated in Study Number Two, Section III, the process observed in calculating the biserial coefficient of correlation is based on some of the properties of the normal probability curve. It was assumed, also, for this section, that the dichotomous traits were normally distributed and continuous, and that the relationship between the two factors was linear.

As in the combined information of the LaDuke and Rostker studies, a biserial coefficient of correlation was calculated for the following qualitative traits:

1. The father was or was not a farmer.

2. The individual did or did not participate in athletics in high school and/or college.

TABLE XV

TEACHER INFORMATION - LINS STUDY (1942)

Teacher No.	Sex	Age	Years of Training	Yrs. Training Beyond High School	Certificate	No. of Schools Taught in	Subject or Grade Taught	Beginning Salary	1947 Salary
2	F	25	4	4	L	1	Science	$1230	$2300
							Home Ec.	1350	2225
3	F	26	4	4	L	3	Home Ec.	1300	2000
4	F	26	4	4	L	1	Home Ec.	1500	2100
6	F	26	4	4	L	3	English	1400	1950
10	F	26	4	4	L	3	Speech	1450	1875
11	F	36	4	4	L	1	Home Ec.	1500	2350
12	F	26	4	4	L	4	Art	$1390.00	$2114.29
					M = 1.71				

TABLE XVI

BISERIAL COEFFICIENTS OF CORRELATION, STANDARD ERROR OF CORRELATION AND LEVEL OF SIGNIFICANCE BETWEEN SIX DICHOTOMIES AND THE CRITERION OF THE FOLLOW-UP STUDY (LINS AND JONES STUDIES COMBINED)

Dichotomy	r E_r	Percent Level of Significance
Father was or was not a farmer	$-.48 \pm .05$	4.88
Did or did not participate in athletics in high school and/or college	$.38 \pm .06$	12.17
Did or did not participate in musical activities in high school and/or college	$.53 \pm .03$	2.86
Did or did not participate in dramatics in high school and/or college	$.00 \pm .24$	
Did or did not participate in community activities as a teacher	$-.33 \pm .07$	17.40
Did or did not belong to a professional organization	$.27 \pm .05$	26.70

N = 18

3. The individual did or did not participate in musical activities in high school and/or college.

4. The individual did or did not participate in dramatics in high school and/or college.

5. The individual did or did not participate in community activities as a teacher.

6. The individual did or did not belong to a professional organization as a teacher.

The standard error of the biserial coefficients of correlation, the biserial coefficients of correlation, and the percent level of significance are found in Table XVI.

The biserial coefficient of correlation between the variable of whether the father was or was not a farmer and the criterion is $-.48 \pm .05$, which is significant at the 4.88 percent level. Selecting as the base that the father was a farmer, the biserial r indicates a negative relationship between success as measured by the criterion of this follow-up study and the fact that their father was a farmer. This correlation would seem to indicate that there was less possibility of their success if their father was a farmer. The negative biserial r in this part of the study between the criterion of this follow-up study and the occupation of the father as a farmer does not agree with the positive correlation found for the group of teachers in the original LaDuke[34] and Rostker[35] studies. This may be due to the fact that the teachers of the LaDuke and Rostker studies were elementary teachers in rural state-graded and village schools, while the teachers of the Jones and Lins studies were secondary school teachers located in urban areas.

A biserial correlation of $-.33 \pm .07$, which is significant at the 17.40 percent level, is found between the criterion and participating in community activities as a teacher. The only significant positive biserial correlation in this part

34. LaDuke, op. cit.

35. Rostker, op. cit.

of the follow-up study is found between the criterion of this follow-up study and participating in musical activities in high school and/or college. The biserial correlation was .53 ± .03, which is significant at the 2.86 percent level. This coefficient of correlation, however, is not in agreement with the findings of previous studies by Durlinger[36], Bossing[37], and Boyce[38].

The other biserial correlations between the criterion and the six qualitative factors fall between .00 ± .22 and .38 ± .06, with none significant at less than the five percent level.

Prediction of Teaching Success - Lins Study

The primary purpose of this section of the follow-up study is to determine if the predictions of teaching efficiency made by Lins in Study Number Two of his thesis were valid and reliable.

The following three factors were among those used by Lins for the prediction of teaching success and were selected for use in the follow-up study:

1. Composite of two interview ratings on general impressions.
2. Rating of ability in assigned activity in Education 73.
3. The final grade for practice teaching in their major.

These three factors were chosen for following the prediction in this study because, through the multiple regression equation, R_O.12 was found to be .56 and also the beta coefficients which indicate the relative weights of the independent variables in the multiple correlation equations were available.

The general form for the multiple regression equation, as given in Section III, Study Number Two, was used.

The proper calculations as indicated by the equation give the following multiple regression equation in raw score form:

$$X_O = .031\,X_1 + .040\,X_2 + .020\,X_3 + 3.103$$

The scores obtained by the teachers, as found by Lins, for the three variables that were employed for prediction in this study, were substituted for the independent variables in the multiple regression equation.

The predicted scores obtained, as contrasted with the actual scores obtained in this follow-up study, are found in Table XVII. For five of the teachers, the predicted score deviates from the actual score, as measured by the criterion of this follow-up study, less than 0.5 of a point, and all of them deviate by less than 1.0.

A further comparison of the predictions is made in Table XVIII where the predicted score is compared with the actual score obtained on the M-Blank in the Lins study. It can be seen that only three of the predicted scores deviate from the actual by less than 0.5 and all scores deviate less than 1.0.

The mean deviation between the predicted score and the actual score in this follow-up study is .44, while the mean deviation between the predicted and actual score in the Lins study is .59. This would seem to indicate that the predictions using the above stated variables were quite reliable and also that the predictions have improved during the several years between the two studies. Although the difference between the original and present rating is not statistically significant, there appears to be some improvement in the teaching efficiency of the participating teachers.

Section IV

SUMMARY AND CONCLUSIONS

This follow-up study has attempted to determine the validity and reliability of the predictions of teaching efficiency made in the original studies and has also attempted to ascertain if any significant changes have occurred in the teachers during the varying number of years between the original studies and the present study.

Since this investigation has made an independent analysis of most of the data for each of the previous studies, the summaries will also be made separately.

LaDuke Study[39]

The information obtained for this part of the follow-up study seems to warrant the following conclusions:

1. The closer relationship between the

38. Durlinger, op. cit.

37. Bossing, op. cit.

38. Boyce, op. cit.

39. LaDuke, op. cit.

TABLE XVII

A COMPARISON OF PREDICTED SCORES OF TEACHING EFFICIENCY USING THREE VARIABLES* AND THE ACTUAL SCORES SHOWING TEACHING SUCCESS AS MEASURED BY THE M-BLANK IN THIS FOLLOW-UP STUDY (LINS STUDY)

Teacher Number**	Predicted Scores	Actual Scores	Difference
1	3.43	4.11	.68
2	3.43	2.94	-.48
3	3.38	3.89	.51
4	3.40	3.39	-.01
5	3.46	3.94	.49
7	3.50	4.28	.78
9	3.41	3.56	.14
11	3.48	4.33	.85
12	3.35	3.33	-.02
			M = .44

*1. Composite of two interview ratings on general impressions.
2. Rating of ability in assigned activity in Education 73.
3. Final grade for practice teaching in major.

** Teachers 6, 8, and 10 omitted because of lack of data.

TABLE XVIII

A COMPARISON OF THE PREDICTED SCORE IN THE LINS STUDY USING THREE VARIABLES AND THE M-BLANK RATING IN THE LINS STUDY

Teacher Number	Predicted Score	Actual Score	Difference
1	3.34	4.25	.82
2	3.43	2.75	-.68
3	3.38	3.50	-.12
4	3.40	3.25	-.15
5	3.46	4.10	.64
7	3.50	4.10	.60
9	3.41	3.75	.34
11	3.48	4.00	.52
12	3.35	1.88	-1.47
			M = .59

measures employed in the original study and the measures employed in the follow-up study after a period of eight years, seems to indicate an improvement in the predictions of teaching efficiency.

2. There appeared to be no significant change in the teaching efficiency of the teachers during the years between the two studies.

3. The M-Blank ratings in this follow-up study are not significantly correlated with the ratings obtained on the supervisory rating scales employed by LaDuke. The correlation coefficients between the criterion of this follow-up study and the ratings by the county superintendents, using the Michigan Rating Scale, Torgerson Rating Scale, and the Almy-Sorenson Scale, were .77, .73, and .81 respectively; the correlation coefficients between the criterion and supervising teachers' ratings, using the above listed rating scales, were .42, .46, and .71, respectively, none of which are significant at the five percent level or less. However, the relationship is closer in the follow-up study than in the original study, which would seem to indicate an improvement in the predictions of teaching success.

4. The correlation between the M-Blank ratings in the follow-up study and the criterion of pupil gain used by LaDuke was .14 ± .44, which is not large enough to be significant.

5. When the information of the LaDuke and Rostker studies is combined, there seems to be a positive relationship between teaching success as measured by the criterion of this follow-up study and the fact that their father was a farmer (r = .36).

6. The correlation coefficient between the criterion of this follow-up study and increase in salary was .09, which is not statistically significant.

7. Participation in community activities shows a negative relationship with the criterion of this part of the follow-up study (r = .48).

Rostker Study[40]

The following conclusions seem warranted for this part of the follow-up study:

1. The closer relationship between the criterion of the original study and the criterion of the follow-up study and the low mean deviation between the predicted and actual scores makes it appear that the predictions of teaching ability made in the original study, after a lapse of eleven years have improved rather than deteriorated.

2. The information obtained does not show a significant change in teaching efficiency during the time between the two studies.

3. The correlation between the criterion of this follow-up study (M-Blank) and the criterion of the original study (pupil gain) is .35 ± .27, which is not large enough so that it cannot be attributed to chance.

4. The correlation coefficients of .27, .19 and .11 between the criterion of this follow-up study and the supervisory ratings in the original study are not large enough to be significant, but they do show a closer relationship between the two measures than the correlations of -.02, .15, and -.01 obtained by Rostker.

5. A rural elementary teacher is more likely to attain success as measured by the criterion if her father is a farmer (biserial r = .36).

6. There is a negative relationship between the teaching success as measured by the criterion and participation in community activities. The correlation coefficient is -.47.

7. A significant correlation coefficient of .85 is found between teaching success as measured by the criterion and the salary received in 1947.

Jones Study[41]

The analysis of the data for this part of the study seems to warrant the following conclusions:

1. Although the difference between the means of the M-Blank ratings in the original study and the M-Blank ratings in the follow-up study is not large enough to be significant, it does indicate a positive gain and a tendency for the teachers to improve in teaching efficiency as measured by the criterion.

2. The predictions of teaching efficiency made by Jones, using the three variables employed in the follow-up study, have im-

40. Rostker, op. cit.

41. Jones, op. cit.

proved after a lapse of four years.

3. The correlation coefficient between the supervisory ratings in the original study and the supervisory ratings in the present study is .48, which is significant at the 2.66 percent level.

4. Of the three other supervisory rating scales employed, only the one used by the Industrial Commission of Wisconsin shows a significant relationship between it and the criterion of the follow-up study.

5. The correlation coefficient of .66 between increase in salary since beginning teaching and the criterion is large enough to be significant at the three percent level.

6. There is not a significant relationship between the criterion of this follow-up study and the salary received in 1947 (r = .38).

Lins Study[42]

For this part of the follow-up study, the following conclusions seem to be warranted:

1. The mean deviation between the predicted scores of teaching efficiency and the actual scores of teaching efficiency in the follow-up study is less than the mean deviations between the predicted and actual scores in the original study. From this, it appears that the predictions of teaching success have improved rather than deteriorated during the four years between the two studies.

2 Although the difference between the means of the criterion of this follow-up study and the criterion of a composite of supervisory ratings in the original study is not large enough to be significant, it seems to indicate a tendency toward improvement in teaching efficiency during the period of four years between the two studies.

3. The correlation coefficient of .73 between the criterion of this follow-up study, which is a rating by one person on the M-Blank, and the composite of five different ratings by five different people on the M-Blank for the original study is large enough to be significant at the 1.64 percent level. This would seem to indicate a significant degree of reliability for the M-Blank.

4. There is a significant correlation between the criterion of this follow-up study and salary received in 1947.

5. A negative biserial r of -.48 between the criterion and whether the teacher's father was or was not a farmer seems to indicate that the teachers of the Lins and Jones studies have greater possibilities of succeeding if their fathers are not farmers.

6. For the teachers of the Lins and Jones studies, there appears to be significant relationship between participation in musical activities and success as measured by the criterion of this follow-up study.

In general it seems that the predictions of teaching efficiency made in each of the four studies, which are a part of this follow-up study, have improved during the lapse of a varying number of years between the original studies and this follow-up study. It also appears that there has been a slight improvement in the teaching efficiency of the groups participating in the original studies, but has not been enough to be statistically significant. Even if the several probabilities are combined, the predictions for the four original studies are not statistically significant.

The relationship between the criterion of pupil gain and supervisory ratings was not significant in any instance and the relationship between supervisory ratings in general, was not significant.

The relationship between other factors considered was not consistent enough to set a pattern.

---

42. Lins, op. cit.

A STUDY OF SOME FACTORS IN PUPIL GROWTH

KENNETH PETER RIESCH*
Mission House College
Plymouth, Wisconsin

SECTION I

The Problem

I.1 Introduction

Educators continue to be interested in factors in pupil growth. Many studies have been made of pupil growth and the attending circumstances. By and large these investigations have been concerned with the construction, measurement, and/or evaluation of single factors contributing to scholastic change. Few reports are concerned with the interplay of several factors.

To be able to predict pupil growth more adequately it would seem that there is a need for multiple factor analysis and the weighting of numerous combined factors. In addition, there must be means of measuring factors affecting this growth either through instruments already refined or more valid and reliable instruments set up. It is necessary, further, to be able to measure adequately the outcomes of growth and to determine what growth or changes are essential to scholastic success.

I.2 Statement of the Problem

The investigation herein reported is a study of certain factors thought to contribute to the social and scholastic growth of pupils. The factors singled out for special consideration are intelligence, past achievement as measured by pre-test scores on measures of social adjustment, personality, conduct, attitude toward the teacher, type of school, and teaching efficiency.

More specifically, this study will attempt to answer the following questions:

1. Are there significant pupil changes in achievement, conduct, attitude toward the teacher, social adjustment, and personality taking place over a three-month period of time?

2. Do factors such as citizenship, intelligence, conduct, teacher attitudes, social adjustment, and personality contribute to scholastic success as measured by gain in achievement?

3. Can gains be attributed to the teacher?

4. Can scholastic success be predicted through a single factor or combination of factors?

5. Can better prediction of academic success be made for rural pupils than for a combined rural-urban group?

6. Are there characteristic differences in attitudes, adjustment, and personality between rural and urban pupils?

7. Are factors such as age, teaching experience, pre-service education, social adjustment, and intelligence associated individually or in combination with teaching efficiency?

8. Can valid predictions of teaching efficiency be made on the basis of the above factors?

9. Are measures of teaching success, i.e., ratings using the M-Blank and a criterion as measured by residual pupil gain, related?

* The author wishes to acknowledge the contributions of various persons who assisted in this study. Grateful appreciation is extended to Dr. C. E. Ragsdale, Professor of Education, for his helpful suggestions in the organization and development of this study, to Dr. M. H. Willing and Dr. T. L. Torgerson, Professors of Education, for their interest and encouragement, to my wife for her assistance in the administration of tests, to Dr. L. J. Lins, director of statistics and research, for his statistical advice, and to the superintendents, teachers, and pupils for their cooperation. Gratitude is extended to my major professor, Dr. A. S. Barr, Professor of Education, to whom the author is deeply indebted for his courtesy, consideration, encouragement, and guidance. His interest and integrity were primary incentives in my beginning and pursuing graduate work.

SECTION II

The Procedure of Investigation

II.1 Design of the Study

An attempt is made in this investigation to determine whether any significant changes took place over a three month period in social adjustment, attitudes, information, and personality of seventh and eighth grade pupils in rural, state graded, and city schools. Various data were collected in 20 school units composed of 22 teachers in 12 rural, six state graded, and two city schools located in Southeastern Wisconsin. The county Superintendent of Schools cooperated with the investigator in the selection of the school units and teachers, used in this study.

The school units were broken down into three divisions which he defined as, (1) rural somewhat isolated, (2) rural not isolated, and (3) state graded with the first division being composed generally of schools located on other than main highways. Since rural schools in Wisconsin often have a limited enrollment, care was taken to select schools having at least a total of six seventh and eighth grade pupils.

Each teacher selected for the research project was notified by the county Superintendent of Schools, and if any objection arose provision was made to make a substitution. In his subjective opinion, each division represented two good, two fair, and two poor teachers.

Upon the suggestion of the writer, the superintendent later filled out the Wisconsin Adaptation M-Blank rating (Appendix A.1)* for each of the rural teachers. This represented a more objective teacher evaluation than he used in his first divisional rating. The superintendents of the two city school systems were requested to fill out the Wisconsin Adaptation M-Blank for the two teachers in his respective seventh and eighth grades.

A week in advance of the testing program, the writer and a rural supervising teacher visited all of the teachers in the 20 school units which were to be incorporated in the study. The purpose of the visit was to answer any questions and to attempt to have the teachers feel that they could make a contribution to education by encouraging their pupils to be cooperative in the testing program.

The testing program was carried out during the second semester of the 1944-45 school year. The investigator and his wife administered the tests. Pre- and post-tests were given to the students with the schools divided so that the same person gave the pre- and post-test for a particular school. This method was employed to help control some of the difficulties that often arise, such as, giving unsatisfactory directions; repeating of the test by another administrator thus giving unsatisfactory results since pupils may not be at ease; lack of interest on the part of administrators; etc.

It was desirable that no student in any school be given a time advantage over any other student in that school. Therefore, the time between all pre-tests and post-tests was exactly twelve school weeks to the hour of beginning administration of the test. That is to say, if a pre-test in a certain field was given at 9 o'clock, the post-test was given at 9 o'clock twelve school weeks later.

The 22 teachers were asked to fill out a sheet known as the Jones' Observational Scale for Rating Citizenship at School (Appendix A.2) for each seventh and eighth grade pupil. A second sheet (Appendix A.3) was also given to the teachers to check any of the 42 items pertaining to Home Environment. The information obtained on both of these sheets was very inadequate, and therefore, after careful study it was decided not to include the results in this investigation.

A study relevant to relationship of home environment and personality adjustment, however was carried out by Stott.[1] From the investigation involving 1,217 children in 14 village schools and 46 one-room country schools, he concluded that "……in general, the way a child feels about himself in relation to the world about him is positively associated with the sort of relationships which exist between himself and his family. There appears to be something about living on a farm near enough to the village or community center and about attending the village school which is relatively favorable to his psychological well being."

II.2 Description of Measures Used

After a review of instruments designed to measure factors of intelligence, social adjustment, attitudes, personality, achievement, and teaching ability, certain tests were selected on the basis of applicability to the study under consideration.

---

* References to Appendices will be found in original thesis on file, University of Wisconsin Library.

1. Leland H. Stott, "Some Environmental Factors in Relation to the Personality Adjustment of Rural Children," Rural Sociology, X (December 1945).

The Wisconsin Adaptation M-Blank as reported previously was used to rate the teachers. The superintendent of each of the city schools rated the seventh and eighth grade teacher in his respective school. Thus a total of 22 teachers were rated.

Measuring instruments used in the study with the frequency of administration appear in Table I. It is noted that wherever possible the same tests were given to students and teachers.

The Cooperative Social Studies Tests (Appendix A.4), devised by Agatha Townsend and Mary Willis, is divided into three parts: I. Fact, Skills, and Applications; II. Terms and Concepts; and III. Comprehension and Interpretation. Scaled scores are provided to enable the administrator to gain directly interpretative information. In using a scaled score a number of 50 corresponds to the expected performance of an "average" individual with three years of study of the subject in grades 7, 8, and 9, tested at the end of the 9th grade. A table of tentative norms based on approximately 5,000 junior high school students was recently published. No report has been made of validity or reliability.

The primary purpose of the Washburne Social-Adjustment Inventory, Thaspic Edition (Appendix A.5), is to determine the degree of social and emotional adjustment of each pupil. The test is arranged so as to give a separate measure of development for six traits which are slightly correlated with intelligence, and more highly correlated with social and emotional adjustment. Three of the traits pertain to emotional adjustments with other people and environment, and the remaining three to self-organization and self-regulation.

The Social-Adjustment Inventory Test has been used and revised over a period of 10 years. The test, for purposes of diagnosis, is divided into eight subtests, one essay, six objective measuring social and emotional adjustment, and one of truthfulness or test accuracy. These are designated by (t) Truthfulness; (h) Happiness; (a) Alienation; (s) Sympathy; (p) Purpose; (i) Impulse-Judgment; (c) Control; and (w) Wishes. The author points out that the initials of the subtests spell the word "Thaspic" omitting the subtest, Wishes.

The Social-Adjustment Inventory is not only said to be a valid instrument, but also a reliable one. The biserial coefficient of validity is reported to be .90. The coefficient of reliability as determined by a retest of approximately 400 college students after an interval of one semester was .92. The probable error for an individual's total social-adjustment score was 5. For the sub-tests, except for the truthfulness score, the approximate range of the coefficients of reliability was from .80 to .90. This information would seem to indicate that the inventory was sufficiently reliable.

In a number of cases the pupils' scores exceeded a truthfulness score of 30. The examiner felt that these pupils could be left in the study as these scores would indicate a very high degree of social maladjustment. Washburne contends that these tests should be discarded because of the low level of truthfulness.

The Right Conduct Test (Appendix A.6) was devised by Harlow A. Wood, Superintendent of Schools, Munising, Michigan. The test is composed of two parts of which Part I consists of 40 short descriptive situations pertaining to questions of social conduct. The pupil is told to select one of the three answers listed under each of the 40 items. Part II is made up of 60 true-false statements relating to questions of social behavior, in which the pupil is told to draw a circle around either true or false.

The purpose of the test is to measure the pupil's social information. At present, coefficients of reliability and validity are not available for the test. Wood has used the test in numerous schools. The results indicate that the instrument has value in determining whether pupils have knowledge of "right" conduct.

Remmer's Scale for Measuring Attitude Toward Any Teacher (Appendix A.7) has been used in several studies with satisfactory results. Tschechtelin, Frances, and Remmers[2] found the scale of value in their study carried out in 1940. The scale consists of 45 statements about the teacher. Pupils are directed to place a plus sign in front of each statement which is true about the teacher, and are assured that their scores will in no way affect their grades in any course.

Scale values are assigned to each of the 45 statements. The final score is merely the scale value of the median item among those that were checked. The measure begins with a positive statement, "is perfect in every way," and ends with negative statements such as the last, "is a failure as a classroom teacher." A scale so constructed is an attempt to supply

2. M. A. Tschechtelin, M. T. Frances, and H. H. Remmers, "Measuring Attitudes of School Children Toward Their Teachers," Journal of Educational Psychology, XXXI (March 1940), pp. 195-203.

TABLE I

TEACHER AND PUPIL TESTS ADMINISTERED WITH FREQUENCY OF ADMINISTRATION

Pupil Tests	Times Given	Teacher Tests	Times Given
Townsend and Willis Cooperative Social Studies Test	2	Townsend and Willis Cooperative Social Studies Test	1
Washburne Social Adjustment Inventory	2	Washburne Social Adjustment Inventory	1
Wood Right Conduct Test	2	Wood Right Conduct Test	1
Remmers Scale for Measuring Attitude Toward Any Teacher	2	Remmers Scale for Measuring Attitude Toward Any Teacher	0
Sullivan, Clark, and Tieg California Test of Mental Maturity	1	Sullivan, Clark, and Tieg California Test of Mental Maturity	1
Clark, Tieg, and Thorpe California Test of Personality	2	Clark, Tieg, and Thorpe California Test of Personality	1

data on the relationship between pupils' attitudes toward their teacher.

The California Test of Mental Maturity, intermediate series (Appendix A.8), and advanced series (Appendix A.9), were administered to the pupils and teachers respectively to determine their intellectual capacity and mental maturity. The test has several significant features such as, (1) a diagnostic evaluation of those mental abilities that are related to success in various types of school activity, (2) a diagnostic profile designed to point out particular abilities, (3) compositie experience of educational leaders in the mental measurement field, and (4) three mental age and intelligence quotients which are, a non-language M.A. and I.Q., a language M.A. and I.Q., and the M.A. and I.Q., characteristic of familiar intelligence.

The battery contains such tests as visual acuity, auditory acuity, and motor coordination. The reason for this organization is to detect pupils with serious defects which might prevent the obtaining of valid diagnosis of their mental maturity. Each of the batteries is built upon the idea of taking into account recognized mental processes or functions.

The various sections of the test show a high reliability. Data was gathered using 400 students in grades 7, 8, and 9, and 100 adults. An analysis was made to determine reliability coefficients computed by the split-halves method and corrected by the Spearman-Brown formula. The results are shown in Table II.

"The probable error of estimate for an individual score varies from three months of mental age when the reliabilities for single grade are .94 and above to five months when the reliabilities are .87 to .89. The probable errors of estimate of individual I.Q.'s for similar reliabilities are 2 and 4 points, respectively."[3]

The California Test of Personality (Appendix A.10) was used to measure the pupils' and teachers' ability to meet social and personal problems in every day living. The major purpose of the test is to discover to what extent the individual is adjusting himself to conditions with which he is confronted. A secondary purpose is to measure whether the individual is developing a normal, happy, and socially effective personality. The test is divided into two parts. Section I deals with self adjustment based upon feelings of personal security, such as, self-reliance, sense of personal worth, and sense of personal freedom from nervous symptoms.

Part II deals with social adjustment based upon feelings of social security such as social standards, social skills, freedom from antisocial tendencies, family relations, and community relations. The two sections aid in determining whether or not the student's basic needs are being satisfied and whether or not he is developing a sense of self-realization and social acceptance. The items included in the personality test were selected on the basis of: judgments of both teachers and principals as to their relative validity and significance, the reactions of students, the extent to which student responses and teacher appraisals agreed, and a study of items using the bi-serial r technique.

The reliability of the California Test of Personality between Section I and Section II is .74. This would seem to indicate that it is desirable to study the student from the standpoint of self and social adjustment since the predictive value is only about 53 percent. Correlations were obtained for 792 cases by the split-halves method corrected by the Spearman-Brown formula. An r of .932 and a probable error of estimated score of 3.7 was found for Section I, Self Adjustment. Section II, Social Adjustment, yielded an r of .873, and a probable error of estimated score of 2.6. For total adjustment the r was .932, and the probable error of estimated score, 3.7.

II.3 Procedure of Scoring Tests

All of the tests used in this study were corrected by the writer. Each test was checked twice in order to eliminate possible errors that may have occurred in correcting them. Several of the tests could have been machine scored, but the writer felt that a better understanding of the problem and data might be possible through hand scoring each particular test.

In scoring the Cooperative Social Studies Test, the essential information needed is printed on scoring keys. Space is provided at the end of each test for entering the necessary figures for computing all raw scores. The number of right answers are entered in the appropriate space, and the amount to be subtracted depends upon the numoer of wrong answers as shown in the chart. All fractions of one-half or less are dropped and fractions greater than one-half increase the amount to be subtracted to the next higher integer. If the resulting difference becomes a negative, the raw score is zero. Upon completing the raw scores, the scaled scores were computed by referring to the conversion tables which are provided.

3. Ibid., p. 2

TABLE II

RELIABILITIES OF THE TOTAL MENTAL FACTORS SCORE AND THE MAJOR TEST SECTIONS*

Variables	Grade			
	7	8	9	Adults
Total Mental Factors	.94	.94	.94	.95
Language Factors	.93	.94	.93	.93
Non-Language Factors	.87	.91	.91	.90
Test A. Memory	.86	.91	.89	.90
Test B. Spatial Reas.	.83	.81	.83	.88
Test C. Logical Reas.	.86	.88	.89	.88
Test D. Numerical Reas.	.91	.92	.92	.94
Test E. Vocabulary	.89	.87	.90	.91
No. of Cases	157	131	112	100
S. D. (M. A. Mo.)	15.0	16.0	17.5	30.0

* Elizabeth T. Sullivan, Willis W. Clark, and Ernest W. Tiegs, California Test of Mental Maturity, Intermediate Series, 1946 Revision. Manual of Directions.

TABLE III

NUMBER OF PUPILS BEGINNING AND FINISHING REQUIREMENTS FOR THIS STUDY DIVIDED ACCORDING TO SCHOOL TYPE

School Type	Total Pupils Beginning	Partly Finished	Total Pupils Completing
Rural Schools	83	12	71
State Graded Schools	116	19	97
City Schools	102	12	90
Total Pupils	301	43	258

The Washburne Social-Adjustment Inventory provides scoring stencils or keys containing a pattern of holes which reveals the incorrect answers of either two or three of the subtests. Each hole in the stencil has a letter designating the subtest to which the item belongs and a number showing the weight to be given to the item if a black mark appears. The score of the "Wishes" is not necessary, since norms are based on the total score made on the objective-type subtests. For this reason the "Wishes" were not used in calculating the final score in this study.

The Right Conduct Test consisted of two parts as previously stated. Part I has a total of 40 items, and Part II a total of 60 true-false statements. The final score is obtained by subtracting the number wrong from 100.

Remmers' Scale for Measuring the Attitude Toward Any Teacher has 45 statements about the teacher. Remmers assigned scale values which are added for those items checked, and the final score is merely the median.

The California Test of Mental Maturity has a key to aid in scoring the test. Each item is considered either right or wrong. Each correct item is marked with a C and if two or more answers are given the answer is considered wrong unless the pupil has attempted to cross out or erase one of the items. The number right is then recorded on the bottom of each page, and transferred later to the first page of the test.

An answer key is used in scoring the California Test of Personality. There is a total of fifteen items in each subsection and the score for each column is the number of student responses which conform with the answer key. Each correct answer is marked with a C. If both YES and NO are marked the answer is considered wrong unless the pupil has attempted to erase or change one of the YES or NO answers. The person doing the scoring must be careful to use the correct column of the answer key for each test column. One adds self adjustment and social adjustment scores to obtain the total adjustment score of each pupil.

II.4 Description of Pupils and Teachers

This investigation began with a total of 301 seventh- and eighth-grade pupils in 12 rural, six state graded, and two city school systems. The number of pupils under consideration was reduced to 258 because of incomplete data; 43 pupils were eliminated because of being unable to comply with all requirements set up, for example, some students moved during the three month lapse of time between the administration of the pre-test and post-test. Other students, who took the pre-test, were absent at the time the post-test was given and therefore were not included in the study. The total number of students in each of the three general school divisions and the number who completed all requirements are presented in Table III.

Twenty-two teachers are included in this study. Their mean age was 35.6 years with the youngest teacher being 20 and the oldest 68. They ranged in experience from one to 43 years with the average being 12.4 while the least amount of pre-service training was one year and the maximum was four years. Average training was 2.3 years.

SECTION III

Report of Findings and Interpretation of Results

This investigation has as its primary purpose the study of various factors which appear to be related individually or in combination to pupil change measured over a three month period of time. The factors under consideration are pupil intelligence, past achievement as given by pre-test scores on measures of social adjustment, personality, conduct, attitude toward the teacher, and teaching efficiency.

III.1 The Measurement of Pupil Gain

Pupil gain or growth was measured with respect to achievement, social adjustment, personality, conduct, and attitude toward the teacher. Pre-tests were given to 301 seventh and eighth-grade pupils in 12 rural, six state graded, and two city school systems. Three months later to the hour a repeat of the tests, hereafter referred to as post-tests, was made to all of the original group available. Because of illness and other factors, 43 pupils did not take all of the post-tests and were, therefore, eliminated, leaving 258. The latter group, then was studied thoroughly.

Each pupil completing the requirements for this study was given the following pre-tests and post-tests: Townsend and Willis Cooperative Social Studies Test, the Washburne Social Adjustment Inventory, the Wood Right Conduct Test, and the Remmers' Scale for Measuring Attitude Toward Any Teacher. In addition, the California Test of Mental Maturity was given. The mean I.Q. as determined from the latter test was 100.69. The standard deviation was 13.24. The means and standard deviations of all other tests are presented in Table IV.

The mean of the group under investigation was 136.6 for the pre-test and 145.3 for the post-test on the California Test of Personality. These means are equivalent to the thirty-fifth and fiftieth percentiles respectively of the group on which the test was standardized. No comparative norms are available for the right conduct test. The 258 pupils stood at about the forty-eighth percentile on the pre-test and the fifty-second percentile on the post-test of the Cooperative Social Studies test when compared with the standardization population. These two tests indicate that the group studied is average in personality and the Cooperative Social Studies.

All mean gains as shown in Table IV were positive with the exception of attitude toward the teacher indicating that some factor or group of factors, operated to bring the teachers into a more unfavorable light during the experimental period. The range of the mean gains was from a positive 8.667 to a -.285. The California Personality had a mean gain of 8.667, the Social Studies Achievement 6.961, the Social Adjustment 3.126, the Right Conduct 2.100, and the Attitude of the Teacher with a -.285.

For purposes of this study, the determination of significance or lack of significance in the differences was measured with respect to the means of the pre- and post-tests utilizing the following equation:

$$\sigma(M_1 - M_2) = \frac{M_1 - M_2}{\sqrt{\sigma M_1^2 + \sigma M_2^2 - 2r_{12}\,\sigma 1\,\sigma 2}}$$

where M_1 is the variance of the post-test, M_2 is the variance of the pre-test, and r_{12} is the Pearson-Product-Moment Coefficient of Rectilinear Correlation between the pre-test and the post-test.

The significance of differences between means is presented in Table V . It appears that significant positive changes have been made in both conduct and achievement since the probability is less than one in 100 that a change of this degree could be attributed to chance alone.

For the other factors, the changes were not statistically significant. There is some reason to believe, under the assumption that the Remmers' Scale for Measuring Attitude Toward any Teacher is a valid measure, that factors operating both in and out of the school situation have not contributed positively but rather have caused a decrease in attitude toward the teacher.

If the Washburne social adjustment inventory alone is considered, there is evidence of pupil maladjustment. On both the pre-test and post-test, the mean of the group falls in the low normal range. Should this test be a true measure of maladjustment, there is indication that individual pupils of the type tested are not well-adjusted since 20 pupils fell in the classification of maladjustment on the pre-test, and 26 fell within this range on the post-test.

From this, one gathers the impression that the influences surrounding the child were not conducive to improvement in social adjustment. Eleven percent of the rural, 10 percent of the state graded, and two percent of the city pupils were shown to be maladjusted on the pre-test, while 13 percent of the rural, 11 percent of the state graded, and six percent of the city were shown to be maladjusted on the post-test. There appears to be no significant differences in attitudes, adjustment, and personality between rural and urban pupils (Appendix B.6).

III.2 Relationship Between Pupil Gain and Factors Thought to Condition Gain

To determine the relationship between the various tests, coefficients of correlation were found for all pre-tests, post-tests, and gains. These with the errors of correlation and the significance level of the correlation are presented in Tables VII, VIII, and IX. The error of correlation was computed using the equation[4]

$$E_r = \frac{1}{\sqrt{N - 1}}$$

This equation was used since it was the desire to determine whether it could reasonably be expected that an r as large as the r obtained would be possible by chance if the true r is zero.

The significance level of the correlation was determined using the following equation:[5] Percent confidence limit 1-2 (area under a normal probability curve) at

$$\frac{r}{\frac{1}{\sqrt{N - 1}}}$$

Referring to Tables VII and XI , one finds that the following pre-tests are significantly interrelated at less than the one percent level,

4. Charles C. Peters and Walter R. Van Voorhis, Statistical Procedures and Their Mathematical Basis (New York: McGraw Hill, 1940), p. 153.

5. L. Joseph Lins, "The Prediction of Teaching Efficiency," Journal of Experimental Education, XV (September 1946), p. 43.

TABLE IV

MEANS AND STANDARD DEVIATIONS OF MEASURES APPLIED TO PUPILS

Measures	M	σ	N
Pre-Tests			
California Personality	136.60	22.76	258
Right Conduct	80.20	12.64	258
Social Studies Achievement	47.10	26.62	258
Social Adjustment	135.21	41.13	258
Attitude of Teacher	8.92	1.32	258
Post-Tests			
California Personality	145.28	24.74	258
Right Conduct	82.30	12.60	258
Social Studies Achievement	54.11	27.48	258
Social Adjustment	137.95	46.32	258
Attitude of Teacher	8.63	2.47	258
Gains			
California Personality	8.68	13.97	258
Right Conduct	2.10	6.83	258
Social Studies Achievement	6.96	11.64	258
Social Adjustment	3.13	12.24	258
Attitude of Teacher	.29	2.48	258

TABLE V

SIGNIFICANCE OF DIFFERENCES BETWEEN MEANS OF PRE-TESTS AND POST-TESTS FOR VARIOUS PUPIL FACTORS

Pupil Factor	Standard Error of Difference
Attitude Toward Teacher	-1.95
California Test of Personality	1.68
Wood Right Conduct	4.38
Cooperative Social Studies	9.94
Washburne Social Adjustment	1.36

TABLE VI

CORRELATIONS BETWEEN GAIN ON PUPIL TESTS AND TEACHING EFFICIENCY (AS MEASURED BY SUPERINTENDENT'S M-BLANK RATING)

Factors	r	E_r	N
Gain Personality	.20	.22	22
Gain Right Conduct	.35	.22	22
Gain Achievement	.22	.22	22
Gain Social Adjustment	.24	.22	22
Gain Attitude	.81*	.22	22

* Significant at one percent level.

personality and right conduct, r = .71 ± .06; personality and social adjustment, r = .65 ± .06; personality and social studies achievement, r = .31 ± .06; right conduct and social adjustment, r = .53 ± .06; right conduct and social studies achievement, r = .71 ± .06; and social adjustment and social studies achievement, r = .43 ± .06. It appears, then, that the tests of personality, social adjustment, and social studies achievement are measuring, to degree, the same thing since they are interrelated. However, they are not related to attitude of the teacher or intelligence as measured here.

What has been stated of the pre-tests is also true of the post-tests with the exception of the fact that the post-test attitude toward the teacher is significantly related to intelligence, r = .16 ± .06, at less than the five percent level. As indicated by Tables VII and XII when compared with Tables VII and XI, there has been change of a greater degree in some factors than in others since the coefficients of correlation, though still significant at less than the one percent level, have changed. The coefficients of correlation are for personality and right conduct, .54 ± .06; personality and social adjustment, .71 ± .06; personality and social studies achievement, .37 ± .06; right conduct and social adjustment, .27 ± .06; right conduct and social studies achievement, .73 ± .06; and social adjustment and social studies achievement, .50 ± .06.

Tables VII and XIII show the general lack of relationship between the pupil gain scores on personality, attitude, intelligence, right conduct, social adjustment, and social studies achievement. The correlation between gain scores for personality and social adjustment is .24 ± .06; for the gains on achievement and social adjustment, .18 ± .06; between pre-test scores on conduct and gain in achievement, r = .21; and between pre-test scores on personality and gain in achievement, r = .20. These are significant at the one percent level.

The composite of gain scores, as outlined earlier, was developed by finding the mean gain on the personality, right conduct, social adjustment, and social studies achievement tests for each pupil. Again referring to Table X, certain significant zero-order coefficients of correlations are found. The following tests when correlated with the composite are significant at the one percent level: California Mental Maturity, .26 ± .06; post-test attitude, .32 ± .06; and pre-test right conduct, .20 ± .06. These coefficients are low and have very little predictive value when individual pupils are considered.

Assuming that the pupil group under consideration is representative of rural pupils, one might conclude that teachers who are rated high by the superintendents when the superintendents use the Wisconsin Adaptation of the M-Blank will conduct their classes in such a way that there would be improvement in the general attitude of the pupil toward the teacher. It also appears from the data previously presented that the pupil tests of personality, social adjustment, and social studies achievement are measuring to degree the same thing. However, these tests and the California Mental Maturity Test are not related.

What has been stated concerning the tests of personality, social adjustment, social studies achievement, and intelligence, is true of the pre-tests as well as the post-tests of the three first mentioned. The relationship between the gains on these three tests and achievement is statistically significant but small. These factors except as previously noted appear not to be related to gains.

III.3 The Prediction of Pupil Achievement

Two measures of scholastic success will be

TABLE VII

COEFFICIENTS OF CORRELATION BETWEEN VARIOUS PUPIL FACTORS EXPRESSED AS Z-SCORES (WITH THE STANDARD ERROR AND SIGNIFICANCE LEVEL OF EACH CORRELATION COEFFICIENT) USING PRE-TEST SCORES

	r	E_r	Significance Level
Personality and Attitude of Teacher	.02	.06	72.64
Personality and California Mental Maturity	.09	.06	14.70
Personality and Right Conduct	.71	.06	.00
Personality and Social Adjustment	.65	.06	.00
Personality and Social Studies Achievement	.31	.06	.00
Attitude of Teacher and California Mental Maturity	.12	.06	5.74
Attitude of Teacher and Right Conduct	.01	.06	94.92
Attitude of Teacher and Social Adjustment	.02	.06	80.26
Attitude of Teacher and Social Studies Achievement	.01	.06	84.14
California Mental Maturity and Right Conduct	.06	.06	35.24
California Mental Maturity and Social Adjustment	.00	.06	95.22
California Mental Maturity and Social Studies Achievement	.09	.06	28.46
Right Conduct and Social Adjustment	.53	.06	.00

Continued.....

TABLE VII (Continued)

	r	E_r	Significance Level
Right Conduct and Social Studies Achievement	.71	.06	.00
Social Adjustment and Social Studies Achievement	.43	.06	.00

TABLE VIII

COEFFICIENTS OF CORRELATION BETWEEN VARIOUS PUPIL FACTORS EXPRESSED AS Z-SCORES (WITH THE STANDARD ERROR AND SIGNIFICANCE LEVEL OF EACH CORRELATION COEFFICIENT) USING POST-TEST SCHORES

	r	E_r	Significance Level
Personality and Attitude of Teacher	.03	.06	63.12
Personality and California Mental Maturity	.00	.06	96.02
Personality and Right Conduct	.54	.06	.00
Personality and Social Adjustment	.71	.06	.00
Personality and Social Studies Achievement	.37	.06	.00
Attitude of Teacher and California Mental Maturity	.16	.06	1.14
Attitude of Teacher and Right Conduct	.03	.06	68.92

considered in this section, namely, (1) social achievement as defined by the gain in the cooperative social studies achievement, and (2) success defined by a composite of gain scores for personality, right conduct, social adjustment, and achievement. The prediction of social studies achievement will be considered first. The intercorrelations for gain in social studies and pre-test scores in this area, conduct, social adjustment, and personality are given in Table XIV.

It is found that the following pre-tests are positively and significantly related to gain in achievement, pre-test conduct, .21 ± .06; pre-test social adjustment, .18 ± .06; and pre-test personality, .20 ± .06. The three coefficients of correlation are found to be significant at the one percent level. Other coefficients with their standard error and significance level are presented in Table VII . The coefficient of correlation between the pre-test on achievement and the gain achievement is no greater than could normally be attributed to chance.

Should one now use the three factors with significant coefficients, the correlation matrix as presented in Table XIV results.

Using the Aitkin Method, a multiple correlation, $R_{0.123} = .230$, is found with the resulting regression equation being $x_0 = .1290x_1 - .0771x_2 - .0578x_3$. If only factors X_1 and X_2 are considered, the multiple correlation, $R_{0.120}$ is .227. The regression equation is $x_0 = .1587x_1 - .0991x_2$. Both coefficients of correlation are significant at less than the one percent level. However, personality does contribute when X_1 and X_2 are used as independent variables. In fact, the pre-test in right conduct is as good, for all practical purposes, for predicting achievement as is the combination of x_1 and x_2 or x_1, x_2, and x_3.

The question might be asked, since both rural and city pupils are considered in this study, whether it might not be more practical to consider the rural and city pupils separately in attempting to discover the relationship between the pre-tests and I.Q. and gain in achievment. There were 90 pupils in the city schools and 168 in the rural.

When the city pupils are considered alone, the coefficients of correlation of Table XIII are found. When the total of 258 pupils were considered, significant correlations were found between gain in achievement and the following pre-tests taken separately: achievement, conduct, and social adjustment. It is noted that when city pupils alone are considered there are no significant correlations at even the five percent level.

However, when one used only the 168 rural pupils, omitting the city pupils from the total group, correlations significant at the one percent level result between the gain in achievement and the pre-test scores on the California Personality, Right Conduct, and Social Adjustment tests. The coefficients of correlation are .29 ± .08, .32 ± .08, and .27 ± .08. These with other coefficients are presented in Table XVI . Again, it may be pointed out that apparently the pre-test in achievement, the attitude toward the teacher, and the intelligence of the pupils do not condition gain in achievement.

It is difficult to explain the apparent difference in ability to predict achievement gain for rural and city pupils. One may account for this, possibly, through considering the outside factors related to the aforementioned gain. It would seem that rural pupils have equal opportunities to a greater degree than city pupils because of the type of work which they do in out-of-school hours. Nearly all of them have "chores" to do on the farm while city pupils have varied experiences and unequal opportunities for study. Some work on paper routes and similar assignments, while others have a great deal of time for home study.

Many people will be interested in pupil gains not merely in subject matter achievement but in conduct, social adjustment, and personality development. Accordingly, a composite of the gains in these several areas was developed for further consideration.

To obtain an overall picture of pupil growth and achievement, a composite of gain scores for personality, right conduct, social adjustment, and achievement was made. The procedure was to add the z-score gains for each pupil, hereafter called the composite gain.

In an attempt to predict this composite gain and ultimately to develop a criterion score of teaching effectiveness based upon residual pupil gains, zero-order coefficients of correlation were obtained between the composite gain and the pre-test, post-test, and gain on the Remmers' Scale for Measuring Attitude Toward any Teacher. The coefficient of correlation was also found between the composite gains and the California Test of Mental Maturity. The Pearson-Product-Moment Coefficients of Correlation were as follows: composite gain and pre-test attitude, .03; composite gain and post-test attitude, .32; composite gain and gain attitude, .02; and composite gain and California Mental Maturity, .26.

Only the correlations between the composite gain and the post-test attitude and California Mental Maturity were significant, both being

TABLE IX

COEFFICIENTS OF CORRELATION BETWEEN VARIOUS PUPIL FACTORS EXPRESSED AS Z-SCORES (WITH THE STANDARD ERROR AND SIGNIFICANCE LEVEL OF EACH CORRELATION COEFFICIENT) USING GAIN SCORES

	r	E_r	Significance Level
Personality and Attitude of Teacher	.07	.06	24.20
Personality and California Mental Maturity	.02	.06	81.26
Personality and Right Conduct	.11	.06	6.88
Personality and Social Adjustment	.24	.06	.00
Personality and Social Studies Achievement	.12	.06	6.00
Attitude of Teacher and California Mental Maturity	.08	.06	18.36
Attitude of Teacher and Right Conduct	-.01	.06	92.04
Attitude of Teacher and Social Adjustment	.05	.06	41.22
Attitude of Teacher and Social Studies Achievement	.02	.06	74.14
California Mental Maturity and Right Conduct	-.01	.06	99.20

Continued.....

TABLE IX (Continued)

	r	E_r	Significance Level
California Mental Maturity and Social Adjustment	.01	.06	93.04
California Mental Maturity and Social Studies Achievement	.01	.06	91.24
Right Conduct and Social Adjustment	.01	.06	90.44
Right Conduct and Social Studies Achievement	.15	.06	1.42
Social Adjustment and Social Studies Achievement	.18	.06	.38
Gain Achievement and Pre-Test Attitude	.05	.06	41.80
Gain Achievement and Post-Test Attitude	.07	.06	25.02
Gain Achievement and Pre-Test Conduct	.21	.06	.08
Gain Achievement and Pre-Test Social Adjustment	.18	.06	.34
Gain Achievement and Pre-Test Personality	.20	.06	.14
Gain Social Adjustment and Attitude of Teachers	.09	.06	16.76
Gain Social Adjustment and California Mental Maturity	.09	.06	17.06
Gain Social Adjustment and Pre-Test Social Adjustment	-.02	.06	31.24
Gain Social Adjustment and Pre-Test Achievement	.06	.06	31.24

TABLE X

COEFFICIENTS OF CORRELATION BETWEEN VARIOUS PUPIL FACTORS EXPRESSED AS Z-SCORES (WITH THE STANDARD ERROR AND SIGNIFICANCE LEVEL OF EACH CORRELATION COEFFICIENT) USING COMPOSITE GAIN SCORES

	r	E_r	Significance Level
Pre-Test California Personality and Composite Gain	.05	.06	41.22
Pre-Test Right Conduct and Composite Gain	.20	.06	.14
Pre-Test Social Studies Achievement and Composite Gain	.03	.06	61.70
Pre-Test Social Adjustment and Composite Gain	-.09	.06	15.56
California Mental Maturity and Composite Gain	.26	.06	.00
Pre-Attitude and Composite Gain	.03	.06	64.56

TABLE XI

SUMMARY OF INTERCORRELATIONS OF PRE-TESTS USING Z-SCORES FOR 258 PUPILS

		x_1	x_2	x_3	x_4	x_5	x_6
Personality	x_1	1.00	.02	.09	.71	.65	.31
Attitude of Teacher	x_2		1.00	.12	.01	.02	.01
California Mental Maturity	x_3			1.00	.06	.00	.07
Right Conduct	x_4				1.00	.53	.71
Social Adjustment	x_5					1.00	.43
Social Studies Achievement	x_6						1.00

TABLE XII

SUMMARY OF INTERCORRELATIONS OF POST-TESTS USING Z-SCORES FOR 258 PUPILS

		x_1	x_2	x_3	x_4	x_5	x_6
Personality	x_1	1.00	.03	.00	.54	.71	.37
Attitude of Teacher	x_2		1.00	.16	.03	.03	.02
California Mental Maturity	x_3			1.00	.06	.04	.06
Right Conduct	x_4				1.00	.27	.73
Social Adjustment	x_5					1.00	.50
Social Studies Achievement	x_6						1.00

significant at the one percent level. Inasmuch as the attitude score is a post-test score and there was only one remaining significant correlation, the regression equation was not set up.

III.4 The Measurement and Prediction of Teaching Success

Of the 22 teachers cooperating in this study, twelve were located in rural, six in state graded, and four in city schools. Factors applied to these teachers were the M-Blank, a rating by the superintendents as outlined previously; the California Test of Mental Maturity; the Washburne Social Adjustment Inventory; and the Cooperative Social Studies Test. The means and standard deviations for these measures are presented in Table XVII . Since a few of the teachers did not take all of the tests, the assumption was made that the most probable value for each teacher would be the mean of all teachers measured. On the basis of this assumption, the mean of each measure was assigned to the teacher who had not taken the particular test.

The superintendents' mean rating, namely, 3.8, of these teachers, was slightly above the theoretical average, 3.0, on the Wisconsin Adaptation of the M-Blank, Cooperative Study of the Secondary School Standards. The scale used was five, outstanding; four, above average; three, average; two, below average; and one, poor. The teaching group stands on the average at the fortieth percentile in comparison with the college freshman I.Q., standard as measured by the California Mental Maturity Test. However, when compared with college graduates, they rank at only about the eighth percentile.

The mean for the teachers on the Washburne Social Adjustment Inventory is at the fiftieth percentile of college students. Eight of the teachers fell in the low normal range while only two teachers fell in the superior range. The amount of pre-service education of the superior range varied from one to four years with a mean of 2.3. The range of teaching experience and teacher age is 1 - 43, and 20 - 68 years, respectively, with mean experience of 12.4 and mean age of 35.6.

The Pearson-Product-Moment Coefficients of Rectilinear Correlation as computed for the teacher factors are presented in Table XVIII. Factors considered were the criterion of teaching efficiency as developed in Section III.1, the Wisconsin Adaptation of the M-Blank, the California Mental Maturity Test, the Cooperative Social Studies Test, the Social Adjustment Inventory, age of the teacher, experience of the teacher, and pre-service education of the teacher.

Associating Tables XVIII and XIX, one notes no significant coefficients of correlation between superintendents' ratings, M-Blank; and the California Mental Maturity; Social Studies Achievement, and Social Adjustment tests. Should the M-Blank ratings be valid measures of teaching efficiency, then it might be concluded that, considering all teachers, the measures used have no apparent value in determining whether or not a teacher is a good teacher.

TABLE XIII

SUMMARY OF INTERCORRELATIONS OF GAINS USING Z-SCORES FOR 258 PUPILS

		x_1	x_2	x_3	x_4	x_5	x_6
Personality	x_1	1.00	.07	.02	.11	.24	.12
Attitude of Teacher	x_2		1.00	.08	.01	.05	.02
California Mental Maturity	x_3			1.00	.01	.01	.01
Right Conduct	x_4				1.00	.01	.15
Social Adjustment	x_5					1.00	.18
Social Studies Achievement	x_6						1.00

TABLE XIV

INTERCORRELATIONS OF GAIN ACHIEVEMENT, PRE-TEST CONDUCT, PRE-TEST SOCIAL ADJUSTMENT, AND PRE-TEST PERSONALITY

Factors		x_0	x_1	x_2	x_3
Gain Achievement	x_0	1.00	.21	.18	.20
Pre-Test Conduct	x_1	.21	1.00	.53	.71
Pre-Test Social Adjustment	x_2	.18	.53	1.00	.65
Pre-Test Personality	x_3	.20	.71	.65	1.00

Since the entire group of 22 teachers may be a more heterogeneous group than the 18 rural teachers when considered alone, it was felt that perhaps teaching efficiency as measured by the M-Blank could be predicted better for the 18 teachers than the total of 22. The coefficients of correla tion are presented in Table XX . Though the coefficients of correlation between the M-Blank and the Achievement and Social Adjustment tests appear to be higher when only the 18 teachers are considered than with the 22, this difference could be attributed to chance since none of the coefficients are significantly different than zero.

What has been stated concerning the relationship between the M-Blank and the various teacher factors can also be said of the relationship between the criterion of teaching efficiency, residual pupil gain, and the various teacher factors. Referring to Table XIX, it is found that none of the coefficients using this criterion are significant. This indicated that one cannot predict teaching efficiency based upon residual pupil gain through the California Mental Maturity, Social Studies Achievement, or Social Adjustment test. There is no association whatever between this criterion and teachers' ages, years of teachers' experience, or years of pre-service education when a group of teachers such as those used in this study are considered.

The M-Blank and the criterion of residual pupil gain are not related by more than an amount that could be attributed to chance, .13 $\pm$ .22. The finding is substantiated by Lins[6] when referring to the Wisconsin Adaptation of the M-Blank and residual pupil gain in his statement that ".....a composite of.....supervisory ratings and pupil gain are not related to greater degree than can be attributed to chance. Whatever, then, is measured by each of the criteria, at least it appears evident that these criteria do not measure the same aspects of teaching efficiency."

One cannot predict either the criterion of residual pupil gain or the M-Blank rating on the basis of the teacher tests used. However, worth investigating is the possibility that supervisors will rate teachers well if they have a superior group of students who "show up" well in the presence of the supervisor. Apparently this is not the case, as shown by Table XXIII, for there are no significant relationships between the superintendent's rating on the M-Blank and the student pre-tests. The respective pre-test factor was developed by finding the mean of all pupils in each particular school and associating that mean with the M-Blank rating of the teacher in that school.

When, however, the relationship is found between the mean pupil factors and the criterion of residual pupil gain, one notes that the pretests of social adjustment and attitude and the I.Q. as measured by the California Mental Maturity test are significantly associated with this criterion. The coefficients of correlation are presented in Table XXII .

Considering the correlation between pupil I.Q. and the criterion of residual pupil gain, one might conclude that the teachers are spending time with the pupils with low I.Q.'s in an attempt to bring them up to the rest of the class in their work. Interpreted differently, since the criterion is based upon gains, it is possible that students who are doing poorly at the beginning of the year will make greater improvement than pupils who rank high at the beginning of the period.

It would seem that if the pupil gains in attitude, personality, right conduct, achievement, and social adjustment could be attributed to the teacher, then there should be a significant correlation or relationship between each respective test gain and the criterion of residual pupil gain. To arrive at this determination, the mean gain for each school on each factor mentioned was found and this was correlated with the residual gain criterion. The resulting coefficients of correlation with their standard errors and significance level are presented in Table XXIII.

The pupil factors do not appear to be significantly related to the criterion at the one per cent level. Gain in achievement is significantly related to teaching efficiency at the five percent level, which indicated that pupils of teachers with a high residual criterion will probably achieve better than those working with a teacher with a low criterion score. Positive but not statistically significant correlations were found between right conduct and attitude and residual pupil gain.

As outlined previously, the superintendents rated the teachers upon the basis of the M-Blank. Their subjective estimate correlated with gain in student attitudes toward the teacher as evidenced by a coefficient of correlation of .81 $\pm$ .22. If then the superintendent rates a teacher high on the M-Blank, there is a high degree of possibility that the children will like that teacher.

III.5 Summary and Conclusions

The central purpose of this section was to

6. Ibid., p. 59.

TABLE XV

CORRELATIONS OF GAIN IN ACHIEVEMENT FOR CITY SCHOOL PUPILS WITH PRE-TESTS SCORES AND I. Q.

Factors	r	E_r	N
California Personality	.00	.11	90
Right Conduct	.04	.11	90
Pre-Test Achievement	-.02	.11	90
Pre-Test Social Adjustment	.10	.11	90
Pre-Test Attitude	.03	.11	90
California Mental Maturity	.04	.11	90

TABLE XVI

CORRELATIONS OF GAIN ACHIEVEMENT FOR RURAL SCHOOL PUPILS WITH PRE-TEST SCORES AND I. Q.

Factors	r	E_r	N
California Personality	.29*	.08	168
Right Conduct	.32*	.08	168
Pre-Test Achievement	.11	.08	168
Pre-Test Social Adjustment	.38*	.08	168
Pre-Test Attitude	.01	.08	168
California Mental Maturity	.13	.08	168

*Significant at the one percent level.

TABLE XVII

MEANS AND STANDARD DEVIATIONS OF MEASURES APPLIED TO TEACHERS

Measures	M	σ	N
M-Blank	3.77	.87	22
California Mental Maturity	106.68	13.89	22
Social Studies Achievement	120.68	18.08	22
Social Adjustment	89.77	28.24	22
Teacher's Age	35.64	20 - 68 Years*	
Teaching Experience	12.43	1 - 43 Years*	
Years of Training	2.27	1 - 4 Years*	

*Range

TABLE XVIII

COEFFICIENTS OF CORRELATION BETWEEN VARIOUS TEACHER FACTORS WITH THE STANDARD ERROR AND SIGNIFICANCE LEVEL OF EACH CORRELATION COEFFICIENT

	r	E_r	Significance Level
Criterion and M-Blank	.13	.22	54.84
Criterion and California Mental Maturity	-.34	.22	12.12
Criterion and Social Studies Achievement	-.39	.22	7.68
Criterion and Social Adjustment	.02	.22	93.62
Criterion and Teacher's Age	-.01	.22	97.60
Criterion and Experience of Teacher	-.00	.22	99.20
Criterion and Years of Training	-.00	.22	98.40
M-Blank and California Mental Maturity	-.10	.22	65.26
M-Blank and Social Studies Achievement	.14	.22	52.86
M-Blank and Social Adjustment	-.00	.22	98.40
M-Blank and Teacher's Age	.14	.22	51.56
M-Blank and Experience of Teacher	.35	.22	10.96
M-Blank and Years of Training	.15	.22	49.64
California Mental Maturity and Social Studies Achievement	.67	.22	.22
California Mental Maturity and Social Adjustment	-.51	.22	1.92
Social Studies Achievement and Social Adjustment	-.40	.22	6.44

TABLE XIX

INTERCORRELATIONS OF TWENTY-TWO TEACHERS USING Z-SCORES

		x_1	x_2	x_3	x_4
M-Blank	x_1	1.00	-.10	.14	-.00
California Mental Maturity	x_2		1.00	.67	-.51
Social Studies Achievement	x_3			1.00	-.40
Social Adjustment	x_4				1.00

TABLE XX

INTERCORRELATIONS OF EIGHTEEN RURAL TEACHERS USING Z-SCORES

	California Mental Maturity	Social Studies Achievement	Social Adjustment	Teacher's Age	Teacher Experience	Years Training
M-Blank	-.01	.27	-.10	.08	.11	.04

TABLE XXI

CORRELATIONS BETWEEN M-BLANK RATINGS AND PUPIL TESTS (MEAN OF EACH SCHOOL), TOTAL OF 22 SCHOOLS CONSIDERED

Factors	r	E_r	N
Pre-Test Personality	.06	.22	22
Pre-Test Right Conduct	.23	.22	22
Pre-Test Achievement	.24	.22	22
Pre-Test Social Adjustment	.24	.22	22
Pre-Test Attitude	-.10	.22	22
California Mental Maturity I. Q.	.25	.22	22

TABLE XXII

CORRELATIONS BETWEEN CRITERION OF RESIDUAL PUPIL GAIN AND PUPIL TESTS (MEAN OF EACH SCHOOL), TOTAL OF 22 SCHOOLS CONSIDERED

Factors	r	E_r	N
Pre-Test Personality	.21	.22	22
Pre-Test Right Conduct	-.01	.22	22
Pre-Test Achievement	-.23	.22	22
Pre-Test Social Adjustment	-.44*	.22	22
Pre-Test Attitude	.48	.22	22
California Mental Maturity I. Q.	-.48*	.22	22

*Significant at five percent level.

TABLE XXIII

CORRELATIONS OF GAIN ON PUPIL TESTS (MEAN OF EACH SCHOOL) WITH CRITERION OF RESIDUAL PUPIL GAIN

Factors	r	E_r	N
Gain Personality	.10	.22	22
Gain Right Conduct	.35	.22	22
Gain Achievement	.53*	.22	22
Gain Social Adjustment	.04	.22	22
Gain Attitude	.28	.22	22

*Significant at five percent level.

study various factors which might be related individually, or in combination, to pupil change measured over a three month period of time. The factors under consideration were I.Q. as determined by the California Mental Maturity Test; pre-test scores on the Cooperative Social Studies Test, Washburne Social Adjustment Inventory, Wood Right Conduct Test, Remmers Scale for Measuring Attitude Toward Any Teacher, the California Test of Personality, and teaching efficiency.

On the basis of this investigation, certain generalizations and conclusions seem warranted:

With respect to the pupils

1. Certain seventh and eighth grade pupils in rural, state graded, and city schools are quite maladjusted. It was found that 10 percent of the state graded, two percent of the city, and 11 percent of the rural pupils were maladjusted at the time of the pre-test as here measured. Factors operating over the three-month period tended to increase this maladjustment so that at the time of the post-tests, 13 percent of the rural, 11 percent of the state graded, and six percent of the city pupils were maladjusted.

2. Significant positive changes in conduct and achievement took place over the three-month period. However, factors operating over this period were either not related to or did not condition significant change in personality, adjustment, or attitude toward the teacher.

3. Gain in achievement is positively and significantly related at the one percent level to the pre-tests of conduct .21 $\pm$ .06; social adjustment, .18 $\pm$ .06; and personality, .20 $\pm$ .06. For all practical purposes, the pre-test of right conduct is as good a measure for predicting gain in achievement as is a composite of the right conduct, social adjustment, and personality pre-tests when weighted in that order: $R_{0.123} = .230$, $R_{0.12} = .227$. The factors of home environment and citizenship were eliminated from the study due to insufficient data.

4. Teaching efficiency measured in terms of residual pupil gain is significantly related (at the five percent level to pupil achievement).

5. There appears to be no significant differences in attitudes, adjustment, and personality between rural and urban pupils.

6. The tests of personality, social adjustment, and social studies achievement are measuring, to degree, the same thing since they are significantly interrelated. However, gains of the various factors are not associated by a greater amount than can be accounted for by chance, with the exception of personality and social adjustment ($r = .24 \pm .06$ - one percent significance level).

7. When the z-score gains of the personality, right conduct, social adjustment, and cooperative social studies tests are composited by finding the mean gain, the following are related at the one percent significance level: California Mental Maturity, .26 $\pm$ .06; post-test attitude, .32 $\pm$ .06; and pre-test right conduct, .20 $\pm$ .16. These coefficients are low and have little individual predictive value.

8. On the basis of the tests used, it is impossible to predict gain in achievement with any acceptable degree of accuracy.

With respect to the teacher

1. There is a wide range in age, teaching experience, and pre-service education of teachers, 20 - 68, and 1 - 4 years, respectively.

2. The tests, Washburne Social Adjustment Inventory, Wood Right Conduct Tests, and California Test of Mental Maturity, administered to the teachers have no apparent value in predicting whether a person will be a good or poor teacher in rural or city seventh or eighth grades. These measures are not significantly related to either the superintendents' general evaluation on the M-Blank or the criterion of residual pupil gain. The same is true of the factors of age, teaching experience, and pre-service education.

3. Whatever is being measured by the M-Blank and the criterion of residual pupil gain, they are not measuring the same thing or, under the assumption that they are measures of teaching efficiency, they measure different aspects of efficiency ($r = .13 \pm .21$).

4. Teachers concentrate upon the maladjusted, low intelligence pupil who enjoys working with the teacher.

5. On the basis of the measures here used it was impossible to predict teaching efficiency with an acceptable degree of accuracy.

BIBLIOGRAPHY

Lins, L. J. "The Prediction of Teaching Efficiency," Journal of Experimental Education, XV (September 1946), pp. 2-60.

Peters, Charles C., VanVoorhis, Walter R., Statistical Procedures and Their Mathematical Bases (New York: McGraw-Hill Book Co., 1940).

Stott, Leland H., "Some Environmental Factors in Relation to the Personality Adjustment of Rural Children," Rural Sociology, X (December 1945).

Tschechtelin, M. A., Frances, M. J., and Remmers, H. H., "Measuring Attitudes of School Children Toward Their Teachers," Journal of Educational Psychology, XXXI (March 1940), pp. 195-293.

A STUDY OF THE DISTRIBUTION OF TEACHER APPROVAL AND DISAP PROVAL AMONG SIXTH-GRADE PUPILS

ALBERT F. deGROAT and GEORGE G. THOMPSON*
Syracuse University

The purpose of this investigation was to develop a scale to measure the social-incentive "climates" which children experience in their day-to-day interactions with teachers. As a second phase of the study an attempt was made to determine some of the psychological characteristics of children living under more or less clear-cut patterns of teacher approval and disapproval. It was hoped that this research undertaking might be profitable in the following ways:

1. To develop one type of evaluation instrument for estimating the relative amounts of teacher approval and disapproval experienced by different pupils.

2. To isolate, in a preliminary way, some of the psychological characteristics of children who are believed by their classmates to be living under different patterns of teacher approval-disapproval.

No attempt has been made in this study to infer any cause-effect relationship that may exist between patterns of teacher approval-disapproval and the psychological growth of children. The results of the present inquiry leave unanswered the question as to whether children have developed certain patterns of behavior partially as a result of teacher approval-disapproval or whether teachers show different evaluative attitudes toward children having particular psychological characteristics. This interesting problem has been left for subsequent investigation in an experiment involving a longitudinal approach of one or more years.

The present research is based on the assumption that the nominating procedure, i.e., having children nominate their classmates to fill certain teacher approval-disapproval descriptions, provides data that would be positively related to patterns of teacher approval-disapproval as determined by an impartial adult observer. The study of this latter relationship has also been left for subsequent investigation. The writers believe that meaningful and worthwhile information is provided by the nominating technique, e.g., consider the case of the child who is mentioned over 100 times as experiencing teacher disapproval and is not nominated at all for teacher approval. Certainly it would appear that a child experiencing this pattern of teacher approval-disapproval would respond differently to the classroom situation than a child who might be experiencing exactly the opposite type of social-incentive conditions. The usefulness of these types of data will, of course, be more adequately realized through a series of investigations, of which the present research is merely an exploratory survey.

Background Literature

To the writers' knowledge the nominating procedure has not been previously applied to the measurement of social-incentive conditions experienced by school children. This general type of psychological scaling technique has been most widely popularized by Moreno (5) and his students in the study of inter-personal relationships. More specifically the particular form of nominating procedure used in the present research was employed in the early Character Inquiry experiments conducted by Hartshorne and May (2). It was from Hartshorne and May's study that the presently employed evaluation instrument acquired its popularly known title, the "Guess Who?" technique. This scaling procedure consists simply of presenting a brief and relatively unstructured description of some behavior that could be attributed to any member of a group and then asking the various members of this group to nominate one or more of their associates whose behavior most nearly fits the description. This "Guess Who?" approach is familiar to most children since it is popularly used in school newspapers, annuals, and joke books. It has been used in recent years by Tryon (9) and Kuhlen and Lee (3) to study the personality characteristics of popular and unpopular children during the adolescent period. The following description is an example of one of the twelve approval items employed in the present research:

"Here is someone who is praised for always having his or her assignments done on

*The writers wish to acknowledge a research grant-in-aid for this study from the School of Education, Syracuse University.

time."

This approach to the measurement of the teacher approval-disapproval atmosphere experienced by pupils has the advantages of being simple to administer and score, of being interesting to the pupils, and of providing a fairly direct evaluation of social-incentive conditions in the classroom as children see them. This latter advantage is important in that it expressed one aspect of the child's reaction to his companions as well as reflecting to some degree teacher approval-disapproval conditions as an impartial adult observer might see them.

The literature devoted to the effects of teacher approval and disapproval on the child's behavior is extensive but somewhat contradictory. The research on the effects of praise and blame as incentives to learning has been carefully evaluated by Schmidt who concluded that, "The preceding review of the literature leaves no definite conclusions to be drawn as to the effectiveness of praise and blame as incentives to learning." (6, p. 7) A more promising approach to the effects of approval and disapproval on learning efficiency was recently reported by Grace (1) in which she attempted to relate personality characteristics to children's responses to approval and disapproval in a learning situation. Although no clear-cut relationships were secured in her study, some promising trends were revealed. The results of Grace's study are similar to those reported by Thompson and Hunnicutt (8) in which they found a relationship between response to praise and blame and the personality characteristics of introversion and extroversion. The relationship between personality characteristics and responses to approval and disapproval needs to be explored more adequately in future investigations. As a working hypothesis the writers believe that children's psychological growth is observably affected by the social-incentive conditions which they typically experience. Since children spend a fairly large portion of their waking hours in a classroom situation where a large share of status rewards are governed by the teacher, it seems important to investigate the patterns of teacher approval-disapproval found in different classrooms. It is also the writers' belief that an understanding and observant teacher will generate social-incentive atmospheres in which approval and disapproval are more or less homogeneously distributed among the various members of her group. The teacher less sensitive to the mental hygiene needs of children will probably be evaluated by her pupils as having "pets" who receive a disapproportionately high amount of commendation and as having less favored pupils who receive the bulk of her disapproval. Hence it may be that this scale of teacher approval-disapproval will be useful in evaluating the success of prospective teachers as well as in estimating the probable behavior of children living under various patterns of social incentives.

Development of the Teacher Approval and Disapproval Scales

The most direct and simple method of developing teacher approval and disapproval scales would be to present only two descriptions to the pupils: "Here is someone who is always being praised by the teacher," and "Here is someone who is always being blamed by the teacher." The disadvantage of such a simple approach is that children would be likely to nominate classmates for approval or disapproval situations on the basis of a general "halo" effect. Another possible approach, and the one adopted for this study, is to sample a fairly large number of commonly occurring situations in which children are commended or rebuked by the teacher for different types of behavior. It was felt that this sampling approach would offset, to some degree, the influence of the "halo" effect and would also help to refresh the child's memory as to some of the situations in which children are praised or blamed by their teachers. The behavior descriptions which were selected are, in the writers' opinion, fairly representative of the types of situations in which children typically receive approval or disapproval from their teachers. They were selected on the basis of controlled observation of teacher-child interactions in several sixth-grade classrooms.

Preliminary Scale

The first teacher approval and disapproval scales that were developed consisted of the following ten approval and ten disapproval items:

Teacher Approval

Here is someone whose work is often pointed out as being very neat.

Here is someone who is often asked by the teacher to be a monitor while she is out of the room.

Here is someone who is often called on to answer questions about the lesson when visitors are present.

Here is someone who is often chosen by the teacher to take leading parts in plays and assembly programs.

Here is someone whose work is often put up on the bulletin board.

Here is someone who is often praised for good writing on all papers handed in.

Here is someone whom the teacher often asks to do errands for her.

Here is someone whom the teacher often lets stay after school to help her.

Here is someone whom the teacher often praises for having good manners.

Here is someone who is often praised for good recitations.

Teacher Disapproval

Here is someone whom the teacher often bawls out for whispering.

Here is someone who is often suspected by the teacher when something happens while she is out of the room.

Here is someone who often has to stay after school to finish the day's work or because of misbehavior.

Here is someone whom the teacher often bawls out for picking on someone else.

Here is someone to whom the teacher often speaks because he or she seldom has the paper or pencil with which to work, or, because he or she is slow in getting them out when it is time to begin to work.

Here is someone whom the teacher often bawls out for shooting paper wads, sailing paper airplanes, tripping people, punching someone, or for disturbing the class in some other way.

Here is someone whom the teacher often scolds for talking without raising his or her hand, or for interrupting others while they are talking.

Here is someone who seldom has his or her work put up on the bulletin board.

Here is someone who is often bawled out by the teacher because he or she pays little attention to what is going on during the class period.

Here is someone who is often scolded by the teacher for handing in untidy papers or because his or her writing is so poor.

These twenty items were "scrambled" and presented with the following instructions to four sixth-grade classrooms in a large urban center:

In this booklet are some word pictures of members of your class. Read each statement and write down the names of the person whom you think the description fits. Remember---

1. Several people may fit one picture. You may write down after each description as many names as you think belong there. (The number of choices was actually limited by the amount of space available for writing--six lines were provided.)

2. The same person may be mentioned for more than one word-picture.

3. If you cannot think of anyone to match a particular word-picture, go on to the next one.

4. Write the given and last names of the persons chosen, not the nicknames or initials.

5. Put you own name down if you think the description fits you.

The results obtained in two typical classrooms are presented graphically in Figure 1. An inspection of the data shows that a few children receive a high percentage of the nominations for teacher disapproval; a similarly small number of children receive a high percentage of the nominations for teacher approval, while the majority of the children receive few, if any, nominations for either teacher approval or teacher disapproval. As discussed later in this report, the relationship between nominations for teacher approval and teacher disapproval appears to vary somewhat from classroom to classroom.

In an effort to determine the internal consistency of the preliminary instruments rank correlation coefficients between the "odd-even" items were computed for both the teacher approval and disapproval scales. These rank correlation coefficients, presented in Table I, demonstrate a fairly high degree of internal consistency among the various items within each of the scales. There appears to be somewhat greater communality among the items making up the disapproval inventory. An analysis of individual items showed that some of the approval and disapproval situations elicited a very small number of nominations. There was an apparent need to delete or revise some of the items, and to add others.

Development of Revised Scales

The revised scales in final form consisted of the following twelve teacher approval items and twelve teacher disapproval items:

Teacher Approval

Here is someone whose work is often put up

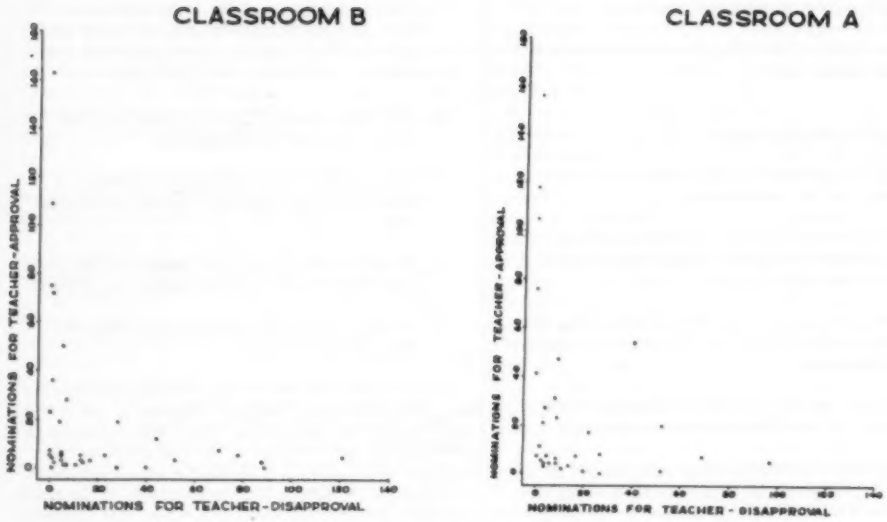

Figure 1

Showing the distribution of teacher approval-disapproval in two sixth-grade classrooms located in a large urban center. Note that a small number of children receive a high percentage of the nominations for teacher approval; also that a small number of children receive a high percentage of the nominations for teacher disapproval. It may also be noted that a fairly large number of children receive few, if any, nominations for either teacher approval or teacher disapproval.

TABLE I

RANK CORRELATION COEFFICIENTS BETWEEN OD-EVEN ITEMS ON THE PRELIMINARY TEACHER APPROVAL AND DISAPPROVAL SCALES*

Classrooms	Number of Pupils	Preliminary Scales	
		Approval	Disapproval
A	31	.67	.90
B	36	.71	.87
C	30	.66	.82
D	20	.83	.63

*Data obtained from four sixth-grade classrooms in a large urban center.

TABLE II

RANK CORRELATION COEFFICIENTS BETWEEN SCORES SCORES OBTAINED BY TEACHER APPROVAL AND DIS-SCALES ON TWO OCCASIONS SEPARATED BY A FIVE-WEEK INTERVAL*

Classrooms	Number of Pupils	Preliminary Scales	
		Approval	Disapproval
I	38	.78	.88
II	29	.81	.86

*Data obtained from two sixth-grade classrooms in a suburban school system.

TABLE III

RANK CORRELATION COEFFICIENTS BETWEEN ODD-EVEN ITEMS ON THE REVISED TEACHER APPROVAL AND DISAPPROVAL SCALES*

Classrooms	Number of Pupils	Revised Scales	
		Approval	Disapproval
1	37	.67	.77
2	33	.72	.78
3	32	.66	.82
4	31	.64	.74

*Data obtained from four sixth-grade classrooms in a small city.

on the bulletin board.

Here is someone whose work is often pointed out as being very neat.

Here is someone who is often praised for good writing on all papers handed in.

Here is someone whom the teacher often chooses to help other children with their work.

Here is someone who is praised for always having their assignments done on time.

Here is someone who has been praised by the teacher because his or her behavior has improved.

Here is someone whom the teacher praises for trying hard, even if the work isn't easy for him or her.

Here is someone who is praised for doing extra work, or for bringing things to school for the class to see.

Here is someone on whom the teacher calls when she wants the right answer.

Here is someone who is praised for helping keep the room clean and orderly.

Here is someone who is chosen by the teacher to take leading parts in plays and assembly programs.

Here is someone whom the teacher often asks to do errands for her or to be a monitor while she is out of the room.

Teacher Disapproval

Here is someone whom the teacher often scolds for whispering.

Here is someone whom the teacher often scolds for disturbing the class in some way (shooting paper wads, chewing gum, etc.).

Here is someone who is often scolded by the teacher because he or she pays little attention to what is going on in class.

Here is someone who has to stay in the room during recess or after school to finish their work.

Here is someone who is often scolded by the teacher for handing in untidy papers.

Here is someone who is often suspected by the teacher when something happens while she is out of the room.

Here is someone whom the teacher often scolds for talking without raising his or her hand.

Here is someone whom the teacher often points out as wasting too much time.

Here is someone who is often pointed out as not doing their best work.

Here is someone whom the teacher often scolds for starting fights or picking on someone else.

Here is someone to whom the teacher often speaks because they do not have their things ready when it is time to begin to work.

Here is someone who sometimes has to go to the principal's office or has to be punished by the teacher.

Six of the original teacher approval items were retained in their exact or slightly revised form and six new items were added. Similarly in the teacher disapproval scale, nine of the original items were retained in their exact or slightly modified form and three new items were added. The instructions for administering the revised scale were similar to those adopted for the preliminary scale. In the revised scale the number of nominations for each item was limited by the amount of space available for writing --only four lines were provided, whereas six lines were available in the preliminary form. It was believed that this revision might "sharpen" the discriminatory power of the scales. Since some overly-conscientious children demonstrate an apparent need to use all of the available space, they may continue to add names which do not necessarily fit the word-description. The only other change in the directions was to instruct the child that he could not nominate himself for any of the situations. These self-nominations, which were permitted in the preliminary form, undoubtedly provide additional and useful information about children, but it appears obvious that they would not be distributed along the same psychological continuum as nominations of fellow classmates.

The revised teacher approval and disapproval scales were administered to two classrooms in a suburban school system on two different occasions, separated by an interval of five weeks. The data obtained from the first administration of the teacher approval and disapproval scales are presented graphically in Figure 2. These data appear to be quite similar to those obtained with the preliminary scales (Fig. 1), in that they show a very small number of children receiving a high percentage of the teacher approval nominations, a few pupils receiving most of the teacher disapproval nom-

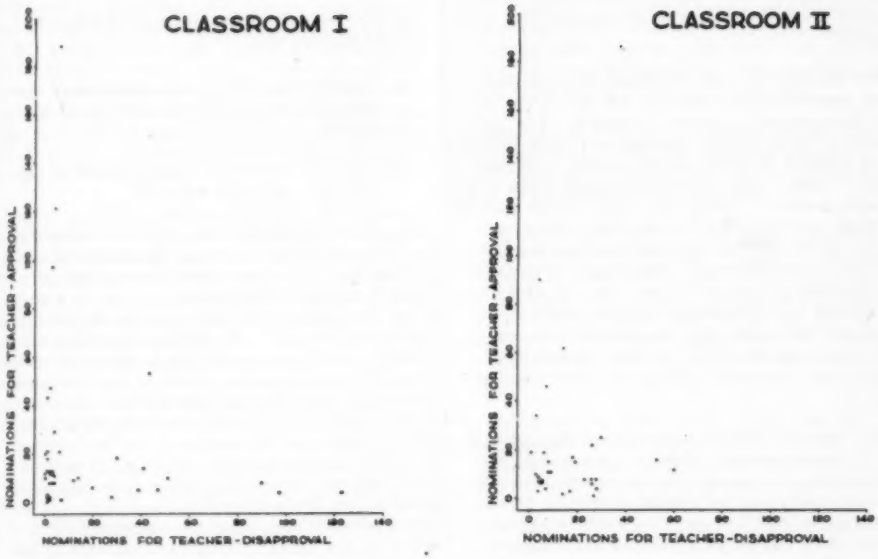
CLASSROOM I
NOMINATIONS FOR TEACHER - APPROVAL
NOMINATIONS FOR TEACHER - DISAPPROVAL
CLASSROOM II
NOMINATIONS FOR TEACHER - APPROVAL
NOMINATIONS FOR TEACHER - DISAPPROVAL

Figure 2

Showing the distribution of teacher approval-disapproval in two sixth-grade classrooms located in a suburban school system. Note that a small number of children receive a high percentage of the nominations for teacher approval; also that a small number of children receive a high percentage of the nominations for teacher disapproval. It may be noted that a fairly large number of children receive few, if any, nominations for either teacher approval or teacher disapproval. The distribution of teacher approval-disapproval in Classroom II, which is a classroom of children of below-average intelligence, is somewhat restricted in total range.

inations, and the majority receiving few, if any, nominations for either teacher approval or disapproval. It is interesting to note that Classroom II, which enrolled children of lower mental ability, reflects the same trend in the distribution of teacher approval-disapproval, but is significantly restricted in range of scores. The conditions that effect this restriction in range of scores may reflect the attitudes of the teacher as well as the mental status of the pupils.

The stability of pupil rankings over the five-week interval with respect to teacher approval and disapproval is shown in Table II. The consistency of the teacher disapproval scale appears to be greater than that of the teacher approval scale. These rank correlation coefficients may be interpreted as indicating a fairly high amount of pupil consistency in nominating their classmates to fill teacher approval and teacher disapproval descriptions. These rank correlation coefficients are fairly satisfactory for tests separated by a five-week interval, since it is probable that some real changes in rank, which would be revealed by a "perfect" measuring instrument, occurred during this five-week interval.

The revised teacher approval and disapproval scales were administered to four sixth-grade classrooms in a small city to provide data for an examination of the individual items. Rank correlation coefficients of the "odd-even" items were computed for the revised teacher approval and teacher disapproval scales. These correlation coefficients, presented in Table III are very similar to those obtained from the preliminary scales. In the revised as well as in the preliminary instruments there appears to be somewhat greater communality among the items comprising the teacher disapproval scale. One might conjecture that it is easier for children to identify individuals receiving teacher disapproval. An alternative hypothesis which one should also consider is that it is easier to construct items which sample the total population of teacher disapproval situations. The revised scales appear fairly satisfactory as to internal consistency of items and test-retest stability.

Psychological Characteristics of Children Experiencing Various Patterns of Teacher Approval and Teacher Disapproval

After developing scales of teacher approval and teacher disapproval that appeared to be fairly satisfactory with respect to internal consistency and test-retest stability, the writers decided to study some of the psychological characteristics of children who might experience the following patterns of teacher approval-disapproval:

1. Children receiving large amounts of approval and little disapproval.
2. Children receiving large amounts of approval and a moderate amount of disapproval.
3. Children receiving little or no approval or disapproval.
4. Children receiving large amounts of disapproval and a moderate amount of approval.
5. Children receiving large amounts of disapproval and little approval.

As a first approximation one can consider these patterns of teacher approval-disapproval as adjoining areas in a two-dimensional plane of teacher approval-disapproval. Thus in a sense, these five patterns constitute a two-dimensional "area-continuum." An individual child's scores on both teacher approval and disapproval would determine his inclusion in one of the five surface patterns. Since the above procedure is merely an approximation adopted for experimental analysis a number of children may not be included in any of the five surface patterns. It was believed by the writers that the pattern approach adopted for this analysis would reveal relationships between psychological characteristics of children and social-incentive conditions in the classroom that a correlational approach with the isolated variables, teacher approval and teacher disapproval, might not reveal.

This phase of the investigation was exploratory in a very real sense. It was designed as an experiment to test some of the more obvious hypotheses which come to mind: (1)Bright children receive more teacher approval than disapproval, whereas dull children experience more disapproval than approval from their teachers; (2) Children who are achieving above the class average receive more approval than disapproval, and vice versa for children below the class average; (3) children who receive little teacher attention, either approval or disapproval, are likely to be withdrawn and retiring. It is possible to develop other hypotheses similar to the ones mentioned above. In this phase of the investigation it was decided to study the intelligence level, the achievement level, and the personality adjustment of children experiencing fairly clear-cut patterns of teacher approval-disapproval. Four classrooms located in a small city in upstate New York were selected for this phase of the investigation. The following tests were administered to the 133 pupils in these four classrooms: (1) Teacher Approval-

Disapproval Scales; (2) California Test of Mental Maturity, Short Form; (3) Progressive Achievement Tests; and (4) California Test of Personality. The means and standard deviations obtained with the standardized tests in each of the four classrooms are presented in Table IV.

The data obtained with the teacher approval and disapproval scales are presented graphically in Figure 3. The data appear to be quite similar to those obtained in the preliminary phases of this study (Figs. 1 and 2) in that they show a few children receiving most of the nominations for teacher approval, a very small number receiving most of the nominations for teacher disapproval, and the bulk of the children receiving few, if any, nominations for either teacher approval or disapproval.

For the present study an attempt was made to isolate those children experiencing each of the five social-incentive patterns listed above. There appeared to be no rational or entirely clear-cut empirical method of isolating these children experiencing different patterns of teacher approval-disapproval. The following rule-of-thumb procedures were finally adopted to isolate within a given classroom those children experiencing the five patterns of teacher approval-disapproval:

1. Those pupils in the upper quartile of teacher approval nominations and falling at or below the median of disapproval nominations were designated as the high approval-low disapproval group (HA-LD).

2. Those pupils in the upper quartile of teacher approval nominations and falling above the median of disapproval nominations were designated as the high approval-medium disapproval group (HA-MD).

3. Those pupils falling at or below the median for both teacher approval and disapproval nominations were designated as the low approval-low disapproval group (LA-LD).

4. Those pupils in the upper quartile of teacher disapproval nominations and falling above the median of approval nominations were designated as the high disapproval-medium approval group (HD-MA).

5. Those pupils in the upper quartile of teacher disapproval nominations and falling at or below the median of approval nominations were designated as the high disapproval-low approval group (HD-LA).

The children included in the HA-LD pattern in each of the four classrooms were combined into a HA-LD group. The same procedure for combining subjects from the four classrooms was adopted for the other four patterns of teacher approval-disapproval. The distributions of teacher approval-disapproval for the five social-incentive patterns are presented graphically in Figure 4. It may be seen that there is a high degree of similarity between some of the points in the HA-LD and the HA-MD groups as well as between some of the points in the HD-MA and the HD-LA groups. If a much larger population of children had been available the rule-of-thumb procedures listed above could have been adjusted so that sharper differentiation between these adjoining groups might have been effected.

A point of interest which certainly needs further investigation is the differential distribution of teacher approval and teacher disapproval nominations in the different classrooms.

In Classroom 1 those four pupils included in the HA-LD classification received 36 percent of the 916 nominations for the twelve teacher approval items. In this same classroom the four pupils included in the HD-LA group received 24 percent of the 780 nominations for the twelve teacher disapproval items.

In Classroom 2 the four pupils in the HA-LD group received 39 percent of the 635 teacher approval nominations. The three pupils in the HD-LA group received 39 percent of the 622 teacher disapproval nominations.

In Classroom 3 the six pupils in the HA-LD group received 64 percent of the 805 teacher approval nominations. The four pupils in the HD-LA group received 37 percent of the 936 teacher disapproval nominations.

In Classroom 4 the seven pupils in the HA-LD group received 72 percent of 886 teacher approval nominations. The five pupils in the HD-LA group received 39 percent of the 1122 teacher disapproval nominations.

It can be seen from these percentages that a very small number of children receive from 36 to 72 percent of the total nominations for approval items. It seems reasonable to believe that this variation in percentages from classroom to classroom may be a meaningful variable for defining the teacher approval-disapproval structure of particular classrooms. It can likewise be seen that a very small number of pupils receive from 24 to 39 percent of the total nominations for disapproval items. This variation in percentages may also provide useful data along the lines indicated above.

It will be noted that only 86 of the total avail-

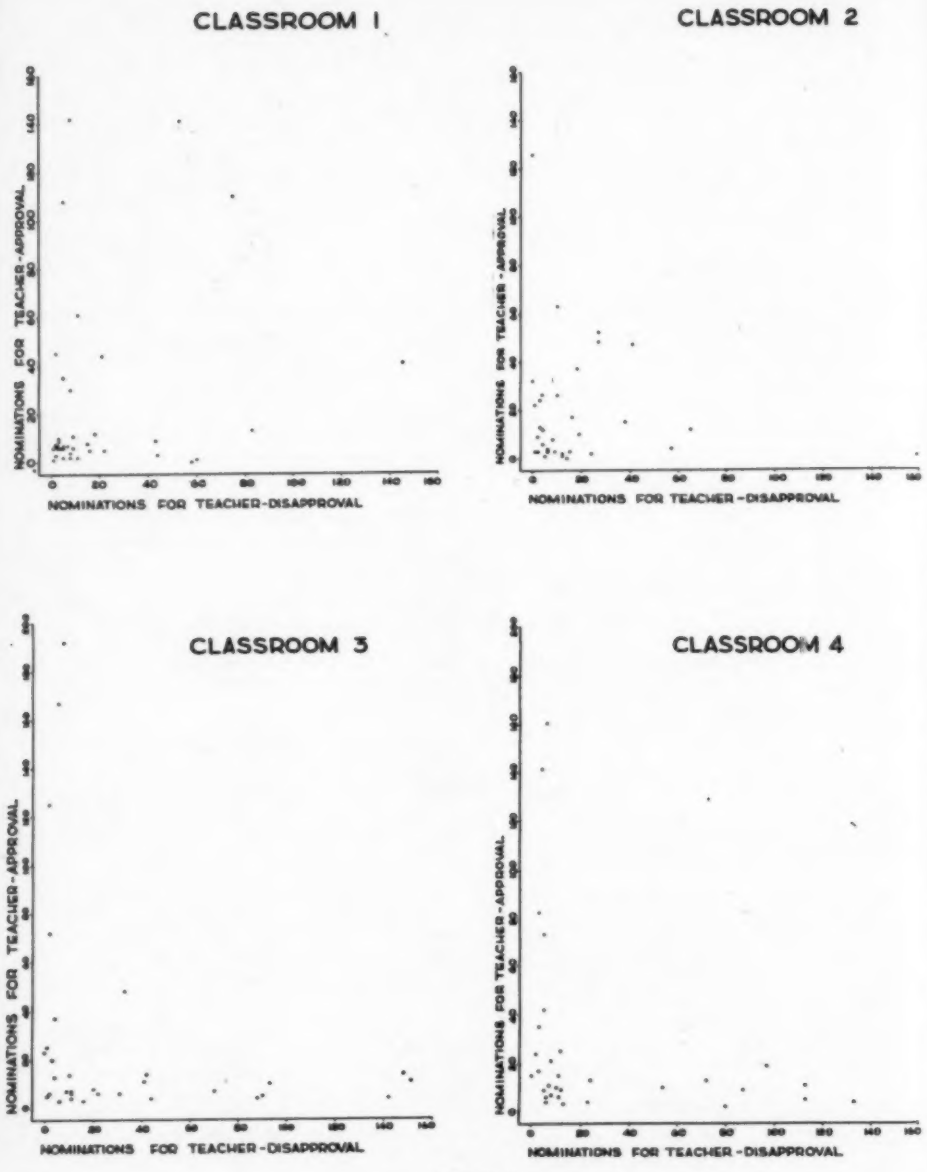

Figure 3

Showimg the distribution of teacher approval-disapproval in four sixth-grade classrooms located in a small city. Note that a small number of children receive a high percentage of the nominations for teacher approval; also that a small number of children receive a high percentage of the nominations for teacher disapproval. It may also be noted that a fairly large number of children receive few, if any, nominations for either teacher approval or teacher disapproval.

TABLE IV

INTELLIGENCE, ACHIEVEMENT AND PERSONALITY TEST SCORES OBTAINED IN FOUR SIXTH-GRADE CLASSROOMS*

Class-rooms	No. of Pupils	Intelligence		Achievement		Personality					
						Self		Social		Total	
		M	S.D.	M	S.D.	M	S.D.	M	S.D.	M	S.D.
1	37	94.14	12.4	255.31	23.6	55.32	8.2	60.59	7.5	115.91	14.3
2	33	90.22	16.3	258.93	32.9	51.32	11.6	60.32	8.0	111.65	18.9
3	31	87.97	19.0	247.06	37.5	49.54	7.1	55.29	9.2	104.82	15.4
4	32	82.75	15.4	242.00	42.6	48.93	10.6	52.44	10.8	101.37	19.9
Total	133	89.08	16.4	251.14	35.0	51.50	10.2	57.45	9.5	108.95	18.3

*Data secured from three schools in a small city in upstate New York.

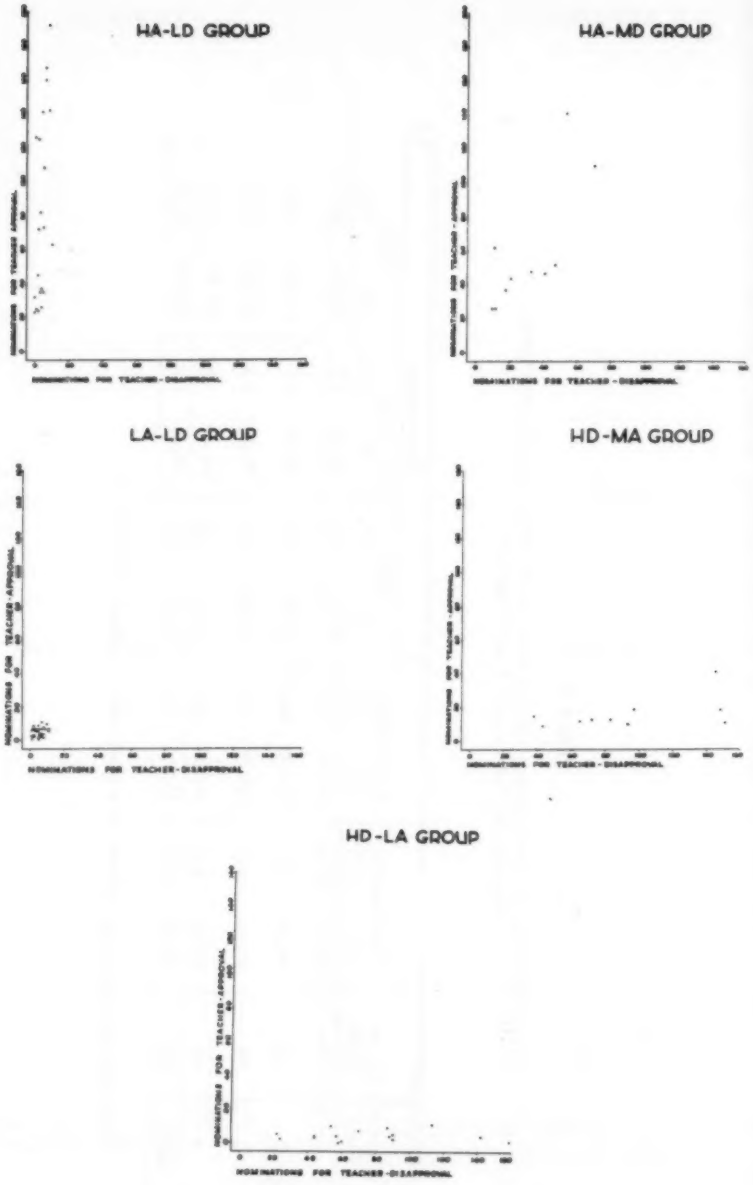

Figure 4

Showing the distribution of teacher approval and disapproval in the five groups selected for experimental comparisons. HA-LD indicates high approval-low disapproval; HA-MD indicates high approval-moderate disapproval; LA-LD indicates low approval-low disapproval; HD-MA indicates high disapproval-moderate approval; HD-LA indicates high disapproval-low approval.

able population of 133 pupils were employed for investigating the psychological characteristics of children experiencing different patterns of teacher approval-disapproval. Some of the cases were omitted because of incomplete test information due to absences, while other cases were discarded because they did not satisfy the selection criteria for any of the five patterns of teacher approval-disapproval.

The means and the standard deviations of the California Mental Maturity, the Progressive Achievement and the California Personality Test scores for each of the five approval-disapproval groups are presented in Table V. It may be seen by inspecting this table that there are some fairly large differences among the standard deviations obtained for the five social-incentive groups on a particular test. One may note in particular the relatively large standard deviation of the HA-LD group on Achievement with respect to achievement scores of the other groups. Because of this apparent lack of homogeneity of variance among the five approval-disapproval groups on certain of the test scores, and because the N's were relatively small it was decided to apply the F-test for homogeneity of variance before testing the significance of differences between means.[1] The results of the F-tests are presented in Table VI.

These results may be interpreted as follows:

1. The five approval-disapproval groups with respect to mental maturity appear to be homogeneous as to variability, i.e. the null hypothesis cannot be rejected for any of the ten comparisons at either the one percent or five percent levels of confidence.

2. The variability of achievement scores appears to be significantly higher for the HD-LA group than for the HA-LD, HA-MD, and LA-LD groups, i.e. the null hypothesis can be rejected at either the one percent or five percent levels.

3. The variability of the total scores obtained on the California Test of Personality appears to be somewhat higher for the LA-LD group and somewhat lower for the HA-LD group, at least in some of the comparisons. The difference in variability between HA-LD and HD-LA is significant at the five percent level. The variability of the LA-LD group is significantly different at the five percent level from both the HA-LD and the HA-MD groups.

4. The variability of the "self" scores of the personality test appears to be relatively homogeneous for the five groups, i.e. the null hypothesis cannot be rejected at either the one percent or five percent levels for any of the comparisons.

5. The variability of the "social" scores on the personality test appears to be somewhat higher for the HD-LA group and somewhat lower for both the HA-LD and HA-MD groups. The variability of the HA-LD group is significantly lower than that of the HD-LA, HD-MA, or LA-LD groups. The variability of the HA-MD group is significantly lower than that of the HD-LA, HD-MA, and LA-LD groups.

The psychological significance of some of these differences in variability is considered in the section on differences between means which follows directly below. The significant differences in variability existing among some of the group comparisons place certain limitations on the interpretation of differences between mean scores, which are presented in Table VII. It will be noted that in this table the results of applying the t-test are presented in parentheses for those comparisons of groups in which there is a significant difference in variability at either the one or five percent levels of confidence. When the scores of two groups are not homogeneous with respect to variability, the interpretation of the results of a t-test cannot be made in a rigorous fashion.[2] The t's presented in Table VII may be interpreted as follows:

1. With respect to mental maturity the HA-LD group is significantly more intelligent than the LA-LD and the HD-LA groups.[3] The HA-MD group is significantly more intelligent than the LA-LD, the HD-MA, and the HD-LA groups. This finding may be interpreted as supporting the hypothesis suggested previously in this report, that the more intelligent children in a typical classroom receive a significantly larger amount of teacher

1. Snedecor, G. W. Statistical Methods (Ames, Iowa: Collegiate Press, 1939), p. 249.

2. Lindquist, E. F. Statistical Analysis in Educational Research (Chicago: Houghton Mifflin, 1940), p. 56.

3. As indicated in the legend to Table VII, only those t's which permit one to reject the null hypothesis at either the one or five percent levels of confidence are considered "significant."

TABLE V

MEANS AND S.D.'S OF MENTAL MATURITY, ACHIEVEMENT, AND PERSONALITY TEST SCORES OBTAINED FOR EACH OF THE FIVE TEACHER APPROVAL-DISAPPROVAL GROUPS

		Patterns of Teacher Approval-Disapproval				
		HA-LD	HA-MD	LA-LD	HD-MA	HD-LA
Mental Maturity	M	97.28	105.10	86.69	87.90	80.12
	SD	14.02	11.62	13.99	12.39	18.03
Academic Achievement	M	274.45	280.10	246.31	247.88	234.43
	SD	22.91	21.64	24.73	22.81	40.98
Total Personality	M	115.95	124.10	108.24	96.11	99.81
	SD	11.47	10.74	19.14	15.09	17.54
Personality "Self-Score"	M	53.86	61.10	50.62	46.11	47.69
	SD	8.27	8.22	11.25	9.65	8.34
Personality "Social-Score"	M	69.10	64.00	57.62	50.00	52.13
	SD	4.57	4.45	8.87	8.06	11.48

Legend:

HA-LD indicates High Approval-Low Disapproval
HA-MD " High Approval-Moderate Disapproval
LA-LD " Low Approval-Low Disapproval
HD-MA " High Disapproval-Moderate Approval
HD-LA " High Disapproval-Low Approval

TABLE VI

VARIANCE RATIOS BETWEEN PAIRED GROUPS USING SNEDECOR'S F-TEST* TO DETERMINE SIGNIFICANCE OF DIFFERENCES IN VARIABILITY**

Mental Maturity				
	HA-MD	LA-LD	HD-MA	HD-LA
HA-LD	1.33	1.02	1.21	1.68
HA-MD		1.31	1.10	1.31
LA-LD			1.19	1.71
HD-MA				2.03

Achievement				
	HA-MD	LA-LD	HD-MA	HD-LA
HA-LD	1.06	1.15	1.08	3.28**
HA-MD		1.22	1.14	3.48**
LA-LD			1.22	2.86**
HD-MA				3.04

Total Personality				
	HA-MD	LA-LD	HD-MA	HD-LA
HA-LD	1.08	2.75*	1.86	2.37*
HA-MD		2.96*	2.00	2.56
LA-LD			1.48	1.16
HD-MA				1.28

Personality "Self" Score				
	HA-MD	LA-LD	HD-MA	HD-LA
HA-LD	1.04	1.82	1.46	1.03
HA-MD		1.75	1.40	1.01
LA-LD			1.25	1.77
HD-MA				1.41

Personality "Social" Score				
	HA-MD	LA-LD	HD-MA	HD-LA
HA-LD	1.01	3.72**	3.34*	6.42**
HA-MD		3.71*	3.32*	6.39**
LA-LD			1.12	1.72
HD-MA				1.92

*See Snedecor (7)

**Those differences which are significant at the five percent level of confidence are indicated by a single asterisk (*), while those which are significant at the one percent level are indicated by a double asterisk(**). The degrees of freedom differ in the several comparisons, hence the ratios cannot be compared directly.

TABLE VII

RESULTS OBTAINED FROM APPLYING FISHER'S "t-test"* TO DETERMINE THE SIGNIFICANCE OF DIFFERENCES BETWEEN THE MEANS OF PAIRED GROUPS**

Mental Maturity				
	HA-MD	LA-LD	HD-MA	HD-LA
HA-LD	1.48	2.59*	1.75	3.17**
HA-MD		3.63**	3.01**	3.74**
LA-LD			.24	1.33
HD-MA				1.15

Achievement				
	HA-MD	LA-LD	HD-MA	HD-LA
HA-LD	.63	1.25	2.68*	(3.52**)
HA-MD		3.74**	2.89*	(3.08**)
LA-LD			.16	(1.15)
HD-MA				.82

Total Personality				
	HA-MD	LA-LD	HD-MA	HD-LA
HA-LD	1.82	(1.61)	3.80**	(3.28**)
HA-MD		(2.42*)	4.44**	3.79**
LA-LD			1.69	1.42
HD-MA				.51

Personality "Self" Score				
	HA-MD	LA-LD	HD-MA	HD-LA
HA-LD	2.21*	1.10	2.16*	2.18*
HA-MD		2.64*	3.46**	3.88**
LA-LD			1.06	.89
HD-MA				.41

Personality "Social" Score				
	HA-MD	LA-LD	HD-MA	HD-LA
HA-LD	2.84**	(6.82**)	(7.94**)	(6.00**)
HA-MD		(2.12*)	(4.49**)	(3.00**)
LA-LD			2.24*	1.74
HD-MA				.47

*See Lindquist (4)

**Those differences which are significant at the five percent level of confidence are indicated by a single asterisk (*), while those which are significant at the one percent level are indicated by a double asterisk (**). The degrees of freedom differ in the several comparisons, hence the ratios cannot be compared directly. Those t's presented in parentheses cannot be interpreted in a rigorous way since they are the result of comparing groups in which there is a significant difference in variability. (See Table VI)

approval, as seen by their classmates, while the less intelligent pupils experience larger amounts of teacher disapproval.

2. In academic achievement the mean score of the HA-LD group is significantly higher than the HD-MA group. The HA-MD group obtained a mean score significantly higher than the LA-LD and HD-MA groups. This may be interpreted as indicating that those children who show evidence of having acquired more subject matter knowledge receive a higher proportion of teacher approval, at least as judged by their classmates.

3. The means of the total scores on the California Test of Personality follow a similar trend. The HA-LD group is significantly "better adjusted" than the HD-MA group. Th HA-MD group shows "superior adjustment" to the HD-MA and the HD-LA groups. At this point it may be well to recall that the total score on the California Test of Personality represents the child's "Adjustment" as he sees it, an "introspectionistic" evaluation. In view of the nature of this test it appears to the writers that the results obtained in this study with respect to teacher approval and teacher disapproval are very provocative. One might expect the traditional teacher to award more approval to the brighter children who do better work. However, it seems extremely doubtful that the average teacher has a great deal of knowledge of, or is in any significant way influenced by, the child's evaluation of his own "adjustment." On the basis of these findings, it appears that the child's evaluation of his own adjustment may be influenced by his level of achievement and/or the teacher's approval-disapproval attitudes toward him. The direction and nature of any cause-effect relationship that may exist between personality adjustment and teacher approval-disapproval remains an interesting problem for future investigation. A developmental study, longitudinal in design, should throw some light on the question of whether those pupils who live in a predominantly teacher disapproval atmosphere become progressively more "poorly adjusted." The present study provides only a hint that this may be the actual state of affairs.

4. The mean "self" scores for the five approval- disapproval groups on the California Test of Personality may be interpreted in a similar way to the total scores presented above. This similar finding might be expected in part since the "self" scores are included as a part of the total scores. As far as "self adjustment" is concerned the HA-LD group is superior to the HD-LA and HD-MA groups. There is one difference in interpretation between the "self" and total scores. The HA-MD group shows significantly better "self adjustment" than any of the other four groups. With the present knowledge available the writers are not in a position to explain why that group, which is highly praised but which also receives a moderate amount of disapproval, shows significantly better "self adjustment" than that group which also receives a high amount of approval and little disapproval. One might hazard a guess that children who receive large amounts of approval and essentially no disapproval may be operating under such intense pressure to conform to social demands as to be actually handicapped in areas of "self adjustment." There is a very slight amount of evidence in the present data to support such a hypothesis in the fact that the HA-LD group receives a higher mean score in "social adjustment" than does the HA-MD group (see conclusion number 5, below). This hypothesis needs to be examined critically by future studies of this type. In conclusion, it may be stated with a high degree of confidence that groups of children who receive high amounts of teacher approval demonstrate "self adjustment" that is superior to groups of children who experience excessive teacher disapproval.

5. As shown in Table V on mean scores, "social" scores on the California Test of Personality decline fairly regularly with progression from the teacher approval to the teacher disapproval end of the "continuum." The differences in the means of the "social" scores are somewhat difficult to interpret in a rigorous way because of the apparent lack of homogeneity of variance among the five approval-disapproval groups. In those groups in which the variability of scores appears to be homogeneous the obtained t's may be interpreted as showing that those groups experiencing large amounts of teacher approval demonstrate superior "social adjustment." Not only do the groups receiving large amounts of approval have higher mean scores in "social adjustment" but they also show significantly less variability in "social adjustment" scores.

As an overall evaluation of the relationship be-

tween teacher approval-disapproval patterns and psychological characteristics of children it may be stated, as a first approximation, that children who are considered by their classmates to be experiencing a high degree of teacher approval coupled with a low or moderate degree of disapproval are more intelligent, are better students, and are better adjusted than their less favored classmates. Whatever cause-effect relationship may exist between teacher approval and superior psychological adjustment cannot be determined in any precise way by the data at hand. However, one gets some hints that children who experience a high degree of teacher approval have a better opinion of themselves, are more "outgoing," and have more confidence in their ability to adjust to social situations.

Discussion

The scales of teacher approval and of teacher disapproval which have been developed in this investigation have interesting possibilities in a number of areas. These instruments should prove helpful in conducting further research of the type described in the second phase of the present study. It is obvious that more work needs to be done to relate the results obtained with these scales to social-incentive conditions in the classroom as an impartial adult judge would evaluate them over a long period of observation. It might also be fruitful to determine the relationship between the results obtained with these scales and social-incentive conditions in the classroom as the teacher sees them. It appears entirely possible that the scales of teacher approval and disapproval might be revised so that they would be more adequate in terms of reliability to measure the social-incentive conditions experienced by individual students. However, it seems desirable to the writers to ascertain the general utility of the scales before attempting further refinement. Some of the possible uses of these scales are as follows:

1. Data of this type might be used to measure one aspect of teacher effectiveness in promoting the general psychological development of pupils. In this way, these instruments would be useful to teacher-preparation programs and to in-service teacher study groups.

2. Data obtained with these scales might provide additional information for predicting the psychological development of children experiencing a certain pattern of teacher approval-disapproval.

The results of the second phase of the present investigation are interesting since they pose many different hypotheses which should be more adequately tested by further research. Some of these hypotheses are as follows:

1. There are individual differences among teachers in generating social-incentive atmospheres of different types. In some classrooms a small number of children receive a disproportionately high percentage of teacher approval. A similar condition exists for teacher disapproval.

2. These results appear to reflect values that are common to our culture, e.g., those children who tend to be more intelligent and to obtain better grades in school receive most of the teacher praise.

3. There appears to be some evidence that children experiencing large amounts of teacher disapproval evaluate their personal worth and general social adjustment as being lower in status than that of their classmates.

4. There is also a hint in the data at hand that the children who receive large amounts of teacher approval and little teacher disapproval may be under considerable social pressure to conform to the conditions for good conduct imposed by the teacher.

It should be noted that these hypotheses, which seem reasonable to the writers, require much additional study before conclusions can be drawn at a reasonably high level of confidence. At the present time a longitudinal study is under way which attempts to throw additional light on some of the problems raised in this investigation.

Summary

The purpose of this investigation was to develop scales of teacher approval and teacher disapproval of the "Guess Who?" type and to determine some of the psychological characteristics of children living under more-or-less clear-cut patterns of teacher approval-disapproval. The scaling procedure devised for this study consisted simply of presenting a brief and relatively unstructured description of some behavior that could be attributed to any member of a group and then asking the various members of this group to nominate one or more of their associates whose behavior most nearly fits the description.

The internal consistency of the items on the preliminary teacher approval and teacher disapproval scales was found to be fairly satisfactory for tests of this type. The stability of the pupil ratings on the revised teacher approval and teacher disapproval scales over a

five-week period was found to be quite adequate for group comparisons--correlation coefficients approximating .80 were obtained. The data obtained with the teacher approval and teacher disapproval scales show that four or five pupils in a typical classroom of 30-35 children receive from 35-70 percent of the total nominations for teacher approval, while four or five different pupils in the same classroom receive 25-40 percent of the nominations for teacher disapproval. The distributions of teacher approval and disapproval for several classrooms are presented in graphic form.

The California Tests of Mental Maturity and Personality and the Progressive Achievement tests were administered to 133 pupils in four classrooms. The children in these classrooms were divided into five more-or-less non-overlapping groups with respect to patterns of teacher approval-disapproval. The results of an analysis of the psychological characteristics of the children included in these five groups are presented in tabular form. Results of various statistical analyses are reported in detail.

As an overall evaluation of the relationship between teacher approval-disapproval patterns and psychological characteristics of children it may be stated, as a first approximation, that children who are considered by their classmates to be experiencing a high degree of teacher approval coupled with a low or moderate degree of disapproval are more intelligent, are better students, and are better adjusted than their less favored classmates. Whatever cause-effect relationship may exist between teacher approval and superior psychological adjustment cannot be determined in any precise way by the data at hand. However, one gets some hints that children who experience a high degree of teacher approval have a better opinion of themselves, are more "outgoing," and have more confidence in their ability to adjust to social situations.

Some of the possible uses in further research and practical application of the teacher approval and teacher disapproval scales are discussed. Several hypotheses about the psychological development and adjustment of children experiencing different patterns of teacher approval and disapproval are presented with suggestions for further research.

BIBLIOGRAPHY

1. Grace, Gloria L. "The Relation of Personality Characteristics and Response to Verbal Approval in a Learning Task," Genetic Psychology Monographs, XXXVII (1948), 73-103.

2. Hartshorne, H., May, M. A., and Miller, J. B. Studies in Service and Self Control (New York: Macmillan, 1929), p. 559.

3. Kuhlen, R. G., and Lee, Beatrice J. "Personality Characteristics and Social Acceptability in Adolescence," Journal of Educational Psychology, XXXIV (1943), 321-340.

4. Lindquist, E. F. Statistical Analysis in Educational Research (Chicago: Houghton Mifflin, 1940), p. 266.

5. Moreno, J. L. Who Shall Survive? A New Approach to the Problem of Human Interrelations (Washington, D. C.: Nervous and Mental Disease Publishing Co., 1934), p. 437.

6. Schmidt, H. O. "The Effect of Praise and Blame as Incentives to Learning," Psychology Monographs, LIII (1941), p. 56.

7. Snedecor, G. W. Statistical Methods (Ames, Iowa: Collegiate Press, 1939), p. 545.

8. Thompson, G. G., and Hunnicutt, C. W. "Effect of Repeated Praise and Blame on the Work Achievement of Introverts and Extroverts," Journal of Educational Psychology, XXXIV (1944), 257-266.

9. Tryon, Caroline C. "Evaluations of Adolescent Personality by Adolescents," Monograph of the Society for Research in Child Development, IV (1939), p. 83.